



Université Libre de Bruxelles



Faculté des Sciences Appliquées
Théorie de l'Information et des
Communications

Quantum Information with Optical Continuous Variables : from Bell Tests to Key Distribution

Promoteur de thèse :

Nicolas CERF

Année Académique 2007–2008

Thèse présentée par

Raúl GARCIA-PATRON SANCHEZ

en vue de l'obtention du grade
de Docteur en Sciences Appliquées

Contents

Preface	i
Acknowledgments	iii
List of Publications	v
Introduction	vii
 I Bell Tests with Continuous Variables	 1
1 Quantum Optics	3
1.1 Quantization of the Electromagnetic Field	3
1.2 Coherent States and Displacements	7
1.3 Linear Optical Operations	9
1.4 Non-linear Optical Operations	15
2 Phase-Space Representation	21
2.1 Introduction to Continuous Variables	21
2.2 Gaussian States	23
2.3 Gaussian Operations	26
2.4 Normal Modes Decomposition	33
2.5 Phase-Space Schmidt Decomposition	35
2.6 Teleportation and Cloning	35
3 Generation of Arbitrary Single-Mode States	43
3.1 Introduction	43
3.2 Generation of a Superposition of $ 0\rangle$ and $ 1\rangle$	45
3.3 Arbitrary Single-Mode State	53
3.4 Efficient State Preparation	58
3.5 Conclusions	61
4 Loophole-free Bell Test	63
4.1 Introduction	63
4.2 Bell Test with Continuous Variables of Light	65
4.3 Feasible Bell Test with Homodyne Detection	67
4.4 Realistic Model	71
4.5 Alternative Schemes	81
4.6 Setup Proposal and Realistic Parameters	84
4.7 Conclusions	86

II	Continuous-Variable Quantum Key Distribution	87
5	Shannon Information Theory	89
5.1	Introduction to Data Compression	89
5.2	Definitions	91
5.3	Entropy Operational Interpretation	96
5.4	Data Merging and Correlation Distillation	100
5.5	Channel Capacity	104
5.6	Decorrelation	107
5.7	Continuous Variables	108
6	Quantum Information Theory	111
6.1	Introduction	111
6.2	Entropies	113
6.3	Quantum Data Compression	117
6.4	Quantum and Classical Correlations	119
6.5	State Merging	121
6.6	Classical Communication	126
6.7	Continuous-Variable Entropy	130
7	Quantum Key Distribution	133
7.1	Introduction	133
7.2	Classical Key Distribution	135
7.3	Quantum Key Distribution	138
7.4	Security Against Eavesdropping	141
7.5	One-Way QKD Protocols	144
7.6	Continuous-Variable Quantum Key Distribution	148
7.7	CV-QKD Entanglement-Based Scheme	153
8	CV-QKD: Individual Attacks	159
8.1	Introduction	159
8.2	Optimality of Gaussian Individual Attacks	160
8.3	Security Analysis using Uncertainty Relations	164
8.4	No Basis Switching Protocol Optimal Attack	173
8.5	Optical Setup Achieving the Optimal Attack	178
8.6	Security Analysis of the Gaussian Protocols	181
9	CV-QKD: Collective Attacks	185
9.1	Introduction	185
9.2	Optimality of Gaussian Collective Attacks	185
9.3	Security Analysis of Gaussian Protocols	187
9.4	Fighting Noise with Noise	196
9.5	Fiber Optic Implementation	203
III	Conclusion and Perspectives	205
IV	Appendices	209
A	The Church of the Larger Hilbert Space	211
B	Partial Measurement of a Bipartite Gaussian State	215

<i>CONTENTS</i>	3
C Properties of $t^{\hat{n}}$	217
D Wigner Representation of Photon Subtraction	219
E Wigner Function from the Fock Basis	223
F Ideal Photon Subtraction	225
G Calculation of G	227
H Incremental Proportionality of Mutual Information	229
I Distance between Purifications	231
J Detail of Calculation of Section (8.4)	233

Preface

When I started writing this dissertation I had two objectives in mind. The first and most important was to present the results that I have obtained during my four years of PhD at the Center for Quantum Information and Communication of the Université Libre de Bruxelles, under the supervision of Nicolas Cerf. The second objective was to write a detailed introduction to the subjects that I have been working on, in order to help new PhD students starting research on these themes.

Since I have been working on two rather different subjects during my PhD, quantum optics with continuous variables and quantum information theory with continuous variables, I have written two independent introductions to fundamental concepts. This explains why nearly half of this dissertation presents fundamental concepts.

Structure of the Dissertation

The dissertation starts with an historical introduction to quantum mechanics, its paradoxes and the new field of quantum information theory. The Introduction situates historically both subjects of my thesis: the Bell tests and quantum key distribution. The rest of the thesis is divided in two parts. Part One concerns different applications of the photon subtraction operation, such as the generation of arbitrary single-mode quantum states of light or the generation of bipartite quantum states of light useful for a loophole-free Bell test. Part Two is centered on the theoretical analysis of the security of quantum key distribution with continuous variables.

Part One starts by introducing in Chapter 1 the fundamental aspects of quantum optics, such as the usual states and operations that can be implemented on a quantum optical table. Chapter 2 revisits quantum optics from the perspective of the phase-space representation, which is the basis of quantum information processing with continuous variables. Chapter 3 presents the concept of photon subtraction as a simple way to generate non-Gaussian states, illustrated by the generation of arbitrary single-mode quantum states of light by combining photon subtractions and displacements. Finally, Chapter 4 presents a proposal of loophole-free Bell test using homodyne detection, where a non-local bipartite state is obtained by a double photon subtraction operation over a two-mode squeezed vacuum state.

Part Two starts with Chapter 5 introducing Shannon information theory in a slightly different way as done in current literature in order to make the transition to quantum information theory easier to the reader, which is presented on Chapter 6. The first part of Chapter 7 is an introduction to quantum key distribution, with the second half presenting the family of continuous variable quantum key distribution protocols based on Gaussian modulation of Gaussian states. Chapter 8 and 9 contain a detailed analysis of the security of those Gaussian protocols: Chapter 8 is centered on individual attacks while Chapter 9 concerns collective attacks.

How to read this Dissertation

An experienced researcher in quantum optics can go directly to Chapters 3 and 4, while an experienced researcher in quantum information with continuous variables can read directly Chapters 8 and 9. Quantum information researchers used to work with discrete variables should read Chapter 2 in order to get acquainted to continuous variables before reading Chapters 8 and 9. Finally a beginner on any of both subjects of this dissertation should read Chapters 1 and 2 for an introduction to quantum optics and phase-space representation, and Chapter 5, 6 and 7 for an introduction to quantum information and quantum key distribution.

Acknowledgments

First of all, I would like to acknowledge my great debt to my advisor Nicolas Cerf for introducing me to the fascinating field of Quantum Information and for his guidance during my PhD. His clever ideas, valuable suggestions and comments were extremely helpful during these four years.

I am also very grateful to him for the atmosphere that he has remarkably succeeded to create at the research group. It was a real pleasure to work with my colleagues during the last four years: Anne-Cécile Muffat, Marie Pinter, Valérie Baijot, Julien Niset, Louis-Philippe Lamoreux, Jérémie Roland, Jaromir Fiurášek, Jonathan Barrett, Evgueni Karpov, David Daems, Kim Nguyen, Koji Maruyama, Stefano Pironio, Sofyan Iblisdir, Serge Massar, Gilles van Assche and Olga López Acevedo.

I would like to thank again Jaromir Fiurášek for his valuable teaching on Quantum Optics and Continuous Variable Quantum Information during my first year of PhD. Without his help, my training period would have probably taken much more time and many of my publications would never have seen the light.

Most of the results of this dissertation come from a fruitful collaboration with the group of Quantum Optics of the Institut d'Optique d'Orsay, I would like to thank the entire group: Philippe Grangier, Rosa Tualle-Broui, Jérôme Wenger, Alexei Ourjoumtsev, together with their collaborators from Thales: Thierry Debuisschert, Jérôme Lodewyck and Simon Fossier. I enjoyed working with them and I hope that in the future we will have more opportunities to collaborate.

I am very fortunate to have had many travel opportunities. Thanks to Zdenek Hradil, Jaromir Fiurášek and Radim Filip for inviting me to the Palacky University of Olomouc, where I enjoyed both working and discussing. I would like also to thank Antonio Acín and Miguél Navascués for inviting me to visit their group at the Institute of Photonic Sciences of Barcelona, where I also had a very fruitful stay.

During the last year of my PhD I had the chance to supervise the work of two students, Loïck Magnin and Jimmy Sudjana. I would like to thank both for their motivation and the excellent work they did.

My gratitude goes to all the people who read a preliminary version of this thesis: David Daems, Loïck Magnin, Julien Niset, Nicolas Cerf and Alba Prieto. I am indebted to Prof. D. Baye for his thorough revision of the manuscript.

Last but not least, I would like to express my gratitude to my parents and family for their support during 27 years, without them it would have been impossible to complete or even start this work. I would also like to thank all my friends, especially those who have accompanied me since my childhood.

List of Publications

Papers in Peer-Reviewed Scientific Journals:

1. R. Garcia-Patron, J. Fiurasek, N. J. Cerf, J. Wenger, R. Tualle-Brouiri, and Ph. Grangier, *Proposal for a loophole-free Bell test using homodyne detection*, Phys. Rev. Lett. 93, 130409 (2004).
2. R. Garcia-Patron, J. Fiurasek, and N. J. Cerf, *Loophole-free test of quantum non-locality using high-efficiency homodyne detectors*, Phys. Rev. A 71, 022105 (2005).
3. J. Fiurasek, R. Garcia-Patron, and N. J. Cerf, *Conditional generation of arbitrary single-mode quantum states of light by repeated photon subtractions*, Phys. Rev. A 72, 033822 (2005).
4. R. Garcia-Patron and N. J. Cerf, *Unconditional Optimality of Gaussian Attacks against Continuous-Variable Quantum Key Distribution*, Phys. Rev. Lett. 97, 190503 (2006).
5. J. Lodewyck, T. Debuisschert, R. Garcia-Patron, R. Tualle-Brouiri, N. J. Cerf, and P. Grangier, *Experimental implementation of non-Gaussian attacks on continuous-variable quantum key distribution system*, Phys. Rev. Lett. 98, 030503 (2007).
6. J. Niset, A. Acin, U. L. Andersen, N. J. Cerf, R. Garcia-Patron, N. Navascues, and M. Sabuncu, *Superiority of Entangled Measurements over All Local Strategies for the Estimation of Product Coherent States*, Phys. Rev. Lett. 98, 260404 (2006).
7. J. Lodewyck, M. Bloch, R. Garcia-Patron, S. Fossier, E. Karpov, E. Diamanti, T. Debuisschert, N. J. Cerf, R. Tualle-Brouiri, S. W. McLaughlin, P. Grangier, *Quantum key distribution over 25 km with an all-fiber continuous-variable system*, Phys. Rev. A (in press), arXiv quant-ph/0706.4255.
8. J. Sudjana, L. Magnin, R. Garcia-Patron, N. J. Cerf, *Heisenberg-limited eavesdropping on the continuous-variable quantum cryptographic protocol with no basis switching is impossible*, arXiv quant-ph/0706.4283 (submitted to PRA).

Conference Proceedings:

1. R. Garcia-Patron, J. Fiurasek, N. J. Cerf, *Proposal for loop-hole free Bell test using homodyne detection*, Proc. of the 7th Int. Conf. on Quantum Communication, Measurement, and Computing, Glasgow, July 24-29, 2004).

Review Chapters:

1. R. Garcia-Patron, J. Fiurasek, N. J. Cerf, *Proposal for loop-hole free Bell test using homodyne detection*, in Quantum Information with Continuous Variables of Atoms and Light. Editors: N.J. Cerf, G. Leuchs and E. Polzik, Imperial College Press, London (2007).

Introduction

Just two dark clouds

The Industrial Revolution marked in the 19th century a major turning point in human social history. This major shift was powered by the developments in two different branches of physics, thermodynamics and electrodynamics, combined with the old Newtonian mechanics. Those three theories could efficiently describe most of the physical phenomena known to date, confirmed by careful detailed experiments for many years.

The general feeling among the physics community at the time was to consider physics as a perfectly assembled and closed structure where there was nearly no possibility for new big discoveries. The following advice by Philipp von Jolly, Planck's professor at Munich, expresses the mood prevalent at that time: "theoretical physics is a noble discipline, [...] but it is improbable that you could add something really new of great relevance" [137]. Physicists were so absolutely certain of their ideas about the nature of matter and radiation that any new concept which contradicted their classical picture would be given little consideration.

Lord Kelvin spoke of only two dark clouds on the Newtonian horizon. The first was the difficulty of conceiving the Earth as moving through ether; the second one concerning the black-body radiation. Since then, physicists have fortunately succeeded to clear the sky of those two dark clouds, but ironically, at the price of changing the horizon. The first was cleared by Einstein's theory of special relativity [66] which changed the "sacred" Newtonian point of view of space and time. The second opened the way to quantum mechanics, a theory that confronts us with seemingly paradoxes which contradict our deepest logical intuition about how nature behaves. Despite the weirdness of the theory, it has succeeded to explain accurately a huge amount of phenomena such as chemistry, nuclear physics and modern electronics which have revolutionized our world.

The triumph of the Quanta

The frequency dependence of the energy emitted by a *black body*¹ was puzzling physicist, since 1859, as they were not able to explain its behavior using the usual tools of thermodynamics and radiation theory. In his 1901 paper, Planck succeeded in deriving the correct shape of the spectrum emitted by a black body, using an assumption that was going to revolutionize physics. He assumed that the energy of the constituting elements of the black body was present in discrete units, *quanta* [150]. Even if Planck thought that his hypothesis was just an artefact that did not refer to real energy exchanges between matter and radiation, there was no return point, quantum theory was born.

¹A black body is an object that is in equilibrium with his environment, absorbing as much radiation as it rejects. These properties make black bodies ideal sources of thermal radiation directly related to their temperature.

In 1905, Einstein took this idea one step further and assumed that the radiation was composed of discrete units of energy, the so called "quanta", which allowed him to explain the photo-electric effect [67], which gave him the 1921 Nobel Prize. This revolutionary idea of *quantization* was applied again in 1907 by Einstein to solid state physics in his seminal paper explaining the specific heat of solids at very low temperature [68]. The first Solvay Conference held in Brussels in the autumn of 1911 "Radiation and the Quanta" heralded the take-off of the new theory, which started to be recognized among the physics community.

At the same time that the theory of radiation was living a revolution on its foundations, so was the theory of matter. In 1909, Ernest Rutherford and his team working in the new research area of radioactivity had the idea to use the massive and positively charged alpha particles to probe the structure of the atom. He observed backward scattering [87] and showed that the atom must have an incredible small but massive center carrying a positive charge. This led him to propose the orbital model of the atom [162], where the recently discovered electron (M. J. Perrin in 1896 [149] and J. J. Thomson in 1897 [186]) orbits around a massive and positively charged nucleus, much like a miniature solar system. However, classical electromagnetic theory predicted that such system should be highly unstable, the electrons radiating their energy until they collapse into the nucleus. Bohr, who had come to Manchester to work with Rutherford, solved the problem in 1913 postulating that the electrons were confined into a discrete set of orbits [25]. In addition, Bohr successfully explained the spectral lines that had been measured for hydrogen and the recently discovered helium [26]. The model improved by Sommerfeld and Pauli set the basis of the nowadays atomic physics, which is at the basis of modern chemistry. The explanation of the ordering of Mendeleev table of elements by this new theory of the atoms [27], is a beautiful example of the astonishing breakthrough realized by Bohr.

Particle or Wave?

By the end of the 19th century most of the physicists were convinced that the light was a wave, as Young's double-slit experiments clearly demonstrated. Surprisingly, Einstein's work on the photo-electric effect needed to assume the light being a particle. The corpuscular behavior was again confirmed by the observation of the Compton effect in 1923 [50], which can be explained as a collision of an electron with a photon. Both incompatible models of light (particle and wave) seemed to be necessary at the same time to explain different experiments, giving birth to the first paradox of quantum physics, the wave-particle duality. At the same time de Broglie postulated that not only light has this wave-particle duality [57], but all existing particles should have a wave behavior, and he succeeded to explain Bohr's model of the atom as stationary waves of the electron. An experimental confirmation of de Broglie revolutionary idea came four years later when the wave behavior of electrons was experimentally demonstrated independently by Davisson [56] and G.P. Thomson ² in 1927 ³.

Despite the important successes of the new theory of quanta, the theory was just a combination of ad hoc quantization rules combined with usual classical physics. A coherent formalism from which one could deduce the ad hoc quantization rules was needed. The answer came in 1925 with two independent developments. The first proposal was Heisenberg's matrix mechanics, which he developed with the help of Max Born and Pascual Jordan [29]. Because Heisenberg's theory was a purely mathematical formalism with no visual aid, physicists of that time preferred the second proposal by

²Ironically, it was the son of J.J. Thomson who demonstrated the particle behavior of the electrons.

³Since then, similar diffractive experiment has been carried out with bigger system, where the actual record is an interference of fullerenes, a molecule consisting of 60 carbon atoms [6].

Schrödinger [168], using a more familiar formalism for the physicist of that time, the wave equation.

Schrödinger developed his famous equation based on de Broglie's concept of matter waves, believing that his approach would reduce quantum mechanics to classical physics by defining his wave function as a density distribution of matter. But his dream vanished when Max Born showed in 1926 that the square of the wave function was indeed the physical probability associated to the particle presence [28]. This new probability was not due to ignorance as in classical thermodynamics but was an intrinsic property of nature. This new probabilistic interpretation combined with the major discovery by Heisenberg of the uncertainty principle in 1927 [99] was a direct attack to the foundations of the "sacred" concept of determinism that had ruled in physics since its beginnings. This tremendous earthquake on the foundations of physics made many physicists dislike this new proposal. Among the most prominent was Einstein, who coined the famous remark "I, at any rate, am convinced that He does not throw dice". This marks the divorce of Einstein with Bohr's and Heisenberg's interpretation of quantum mechanics, called the Copenhagen interpretation. Einstein's disagreement with the "orthodox interpretation" of quantum mechanics will last until his death in 1955.

The Interpretation Paradoxes

Although the Copenhagen interpretation provided a strikingly successful calculation recipe, many physicists continued to think that quantum mechanics should be described by a more fundamental and deterministic physical theory. Another of those unhappy physicists was Schrödinger, who in 1935 proposed his famous "cat's thought experiment" [169] which attempts to illustrate the incompleteness of Bohr's interpretation of quantum mechanics when going from subatomic to macroscopic systems. Schrödinger imagined a cat placed in a box with a radiative atom and a Geiger counter that activates a deadly poison fume killing the cat if the radiative atom decays. The radiative atom being in a superposition of decaying and not decaying, it translates into a macroscopic superposition of a dead and alive cat, which seems not acceptable as a superposition of macroscopic object has never been observed. The response of the Copenhagen interpretation to this paradox was that the act of observation collapses the wave function in one of both states of the cat.

Even if most physicists were more concerned with practical applications of quantum mechanics than by its interpretation problems, Einstein did not give up. After arriving to Princeton escaping from the Nazi rise to power, he proposed together with Boris Podolsky and Nathan Rosen another "thought experiment" now called the EPR paradox [69] after its authors names. The experiment considers a very special state composed of two spatially separated particles that interacted in the past, having stronger correlations than any classical system can have. Schrödinger coined the term *entanglement* for this property. The authors suggested in their work a contradiction between quantum mechanics and three assumptions that they considered to be necessary in any reasonable physical theory: (i) causality; (ii) a definition of reality of a physical quantity; (iii) separate systems should maintain separate identities. The authors being convinced that quantum mechanics was then an incomplete theory, they suggested that the probabilistic structure of quantum mechanics should be described by an underlying deterministic substructure.

Einstein and Schrödinger were not the only ones to disliked the special role of measurement in Bohr's interpretation and the resulting probabilistic structure. In 1957, Everett proposed in his PhD dissertation a new interpretation of quantum mechanics which eliminates the special role that measurement has in the Copenhagen interpre-

tation [74]. The crucial idea is that he considered the measurement process as an interaction between the object and the measurement apparatus ruled by quantum mechanics, governed by the same rules as any interaction between two quantum system. But the price to pay for solving the collapse problem is extremely high, as an even stronger interpretation problem appears. Everett's interpretation allows every possible outcome of each event to exist in its own "history", the superposition of "Schrödinger cat" becoming a reality!. The apparent randomness being just a perception of the observer inside each "history". Everett's interpretation became known as the "many worlds", which has been an endless source of inspiration for science fiction writers.

Rather than remain stuck by interpretation problems of the theory, most physicists took a pragmatic approach and continued using quantum mechanics to study unsolved problems. The reason was simple but astonishing, quantum mechanics powered by the new mathematically rigorous formulation developed by Paul Dirac [60]⁴ and John von Neumann [192] was succeeding to explain an incredible amount of phenomena such as the basis of chemical reactions, nuclear physics and predicting the existence of new particles such as the positron or neutron. The success of quantum mechanics was so important that discussing about the interpretation and paradoxes was nearly a taboo, a domain relayed to philosophers of science.

Entanglement Becomes a Reality

The EPR argument gained a renewed attention in 1964, when John Bell, a researcher at the European accelerator laboratory CERN, derived his famous inequalities [15]. Bell showed that any deterministic substructure model following Einstein's conception of Nature (also called an *hidden-variable model*), while it satisfies causality, it yields predictions that significantly differ from those of quantum mechanics. The merit of Bell inequalities lies in the possibility to test them experimentally, allowing physicists to test whether either quantum mechanics or "hidden-variable models" is the correct description of Nature.

A decade later, an important technology development allowed researchers to implement sources of entangled states, allowing for the first time to test the foundation principles of quantum mechanics. The first Bell test was carried out by Freedman and Clauser in 1972 [80], which was later improved by Aspect, Grangier, Dalibard and Roger, who performed experiments at the beginning of the 80's [8, 9, 10] and more recently by Zeilinger's team in 1998 [195]. All the performed experiments observe the violation of Bell inequalities as predicted by quantum mechanics. But from a logical point of view, these experiments do not succeed in ruling out a "hidden-variable model", as an extra assumption is necessary: the pairs of photons registered by the detectors form a fair sample of the emitted pairs.

Even if there remains some controversy about the interpretation of the results of Bell experiments⁵, most of the physicists agree that Nature does not behave as Einstein's model of "hidden variable" predicts. The confidence in quantum mechanics is strengthened by two other experiments, the *delayed choice experiment* [110, 116] and the *GHZ paradox* experiment [143], which cannot be explained by any classical model. But from a purely logical point of view "hidden-variable models" have not been ruled out yet, as a loophole-free Bell test has not been carried out yet. In the first part of this dissertation, we propose an experimentally feasible setup capable of carrying such a Bell test in the near future.

In the last two decades quantum physics has lived a "renaissance" thanks to a

⁴Interestingly, Dirac proved that Heisenberg and Schrödinger interpretations were equivalent.

⁵For a detailed description of this controversy see the introduction of Chapter 4 of this thesis where we present an experiment that if implemented would definitively close the debate.

technological breakthrough that has allowed experimentalists to manipulate for the first time individual particles and generate entangled states. Examples of such revolutionary technologies are: the cooling down to absolute zero temperature, which has allowed to trap individual atoms and ions; the generation of pairs of entangled photons; the generation of Bose-Einstein condensates; or the superconducting quantum dots. This has allowed physicists to transform the old "thought experiments" into real experiments which test the foundations of quantum mechanics every day in the laboratories, with no failure of quantum theory predictions observed up to date.

The possibility of addressing atoms and photons individually has recently raised the following question "What happens if we encode information into microscopic particles?". The answer is a promising new field of physics called Quantum Information. This field results from merging quantum mechanics with two of the most important theories of the 20th century, information theory and computer science, which developed after the seminal works of Claude Shannon [175] and Alan Turing [187], respectively.

Information is Physical

Before continuing the presentation of quantum information let us jump back to the 40's. The Second World War acted as catalyzer for the research and development of new technologies. Even if most of the developments took place in the domain of engineering (rockets, jet propulsion or nuclear weapon research), also new branches of applied mathematics, such as *operational research*, were born during the war. But probably the two most important examples of theories born at that time are, information and communication theory and computer science, as they are at the origin of another technological revolution comparable to the Industrial Revolution, which has changed our way of living in the last decades.

The landmark event that established the discipline of information theory, was the publication of Claude E. Shannon's classic paper "A Mathematical Theory of Communication" in 1948 [175], kept secret during the war. Shannon information science has become since 1948 a flourishing field, with numerous applications such as error correcting codes used in digital communication and data storage. During the first years of the development of information theory, the information was seen as an abstract mathematical concept. This view smoothly switched to a more physical description of information, where one can see a storing media (e.g. hard disk) as a collection of information units that can be in two different physical states, encoding logical zeros and ones. This new physical point of view on information came along with some interesting theoretical results such as Landauer's principle [124] linking the erasure of information to the dissipation of energy and the concept of reversible computation.

Let us return to the 90's. During the last decade physicists started to study the effects of encoding information in quantum objects, giving birth to Quantum Information. This new field offers novel applications such as quantum computation and quantum cryptography, which are impossible to get using classical information encoded in macroscopic objects, as done in current IT applications. Quantum computers are a promising technology that would allow decreasing the calculation time of many interesting problems such as factoring large numbers or searching an unsorted database, which are used in numerous applications. Quantum cryptography is the most developed application of Quantum Information, enabling two distant partners linked by a quantum channel and a usual communication line to distribute a secret key unknown to a potential eavesdropper. In contrast to nowadays cryptographic protocols, such as RSA [160], which are based on the difficulty of solving some mathematical problems such as factorization of large numbers (that could be broken by a quantum computer), the security of quantum key distribution is stronger since it is assured by the laws of

physics. During recent years, quantum cryptography has been the object of a strong activity and rapid progress, and is now extending its activity into commercial products proposed by some startups.

The interest of Quantum Information not only resides in its applications, the theory is also interesting in itself as it gives a new insight on the foundations of quantum mechanics, which can be experimentally tested thanks to new technological developments. This new way of thinking about quantum mechanics is starting to influence other fields of physics where quantum effects are present, such as solid state physics, resulting in unexpected and interesting contributions to those fields.

Continuous Variables

As for classical information, the quantum information can be divided into two families depending on the encoding techniques: discrete variables (quantum-bit) and continuous variables. Since the experimental demonstration in 1998 of unconditional quantum teleportation [83], continuous variable quantum information has become a flourishing field with many practical advantages over its quantum-bit counterpart, especially for protocols related to communication. For example, if one uses the quadratures of the electromagnetic field to encode continuous variables, linear optical circuits, coherent detection and feed-forward are enough to implement many interesting protocols such as cloning and quantum key distribution. Together with the simplification of the processing operations, the use of coherent detection reaches a much higher optical data rate than with usual photodiodes, allowing faster, cheaper and more efficient detectors.

Entanglement being the key resource for many quantum information applications, the generation and manipulation of continuous-variable entangled states has been a very important field of research during the last years, both theoretically and experimentally. The usual way of generating continuous-variable entanglement in experiments, such as in the teleportation experiment, is based on parametric amplification processes. Interestingly, continuous-variable experiments do not suffer from two drawbacks present in qubit-based implementations: (i) current optical sources of entangled qubits do not succeed generating entanglement on demand; (ii) the measurement in the basis of entangled states is not unconditional. The easiness of the generation, manipulation and measurement of entangled states makes continuous-variable quantum information even more interesting.

The large majority of quantum states of light that are currently accessible in quantum optics labs are the so-called Gaussian states, presented in Chapters 1 and 2 of this dissertation. The major part of the optical operations that are nowadays accessible are called Gaussian operations, as they preserve the Gaussian property of the states. Gaussian states and Gaussian operation are crucial tools for continuous-variable quantum information which have been extensively studied during the last years.

Unfortunately it was recently discovered that not all quantum information applications are implementable using just Gaussian states and Gaussian operations. For example, universal quantum computation and entanglement distillation need more sophisticated non-Gaussian operations, which has increased the interest over non-Gaussian operations in the last years. One of the simplest non-Gaussian operations being accessible today is the photon subtraction operation [139], which is the core element of the first part of this PhD dissertation. In Chapter 3 we propose a technique to generate highly non-Gaussian single-mode states of light based on this novel operation, Chapter 4 proposes an experimental setup capable of realizing a loophole-free Bell test using the photon subtraction operation.

In the second part of this PhD dissertation, we study a continuous-variable version of quantum key distribution. Chapter 8 and Chapter 9 treat a complete security

analysis of a family of protocols based on Gaussian modulation of Gaussian states. Those protocols are very interesting due to the practical advantages of continuous variables described above.

Part I

Bell Tests with Continuous Variables

Chapter 1

Quantum Optics

1.1 Quantization of the Electromagnetic Field

In this section we introduce the quantization of the electromagnetic field and different state representations such as the Fock state and quadrature basis.

Classical Description of Free Electromagnetic Field

The Maxwell equations of the electromagnetic field in the vacuum relate the electric \mathbf{E} to the magnetic field \mathbf{H} . Since there is no current or electric charge present in the vacuum, the equations have the form:

$$\nabla \times \mathbf{E} = -\epsilon_0 \frac{\partial \mathbf{H}}{\partial t}, \quad (1.1)$$

$$\nabla \times \mathbf{H} = \mu_0 \frac{\partial \mathbf{E}}{\partial t}, \quad (1.2)$$

$$\nabla \cdot \mathbf{E} = 0, \quad (1.3)$$

$$\nabla \cdot \mathbf{H} = 0, \quad (1.4)$$

where ϵ_0 and μ_0 are the free space permittivity and permeability, respectively, satisfying $\mu_0 \epsilon_0 = c^2$ where c is the speed of light in vacuum. It follows that $\mathbf{E}(\mathbf{r}, t)$ satisfies the wave equation

$$\nabla^2 \mathbf{E} - \frac{\partial^2 \mathbf{E}}{\partial t^2} = 0. \quad (1.5)$$

This equation has a solution in terms of forward (backward) propagating plane waves, traveling at the speed of light c ,

$$\mathbf{E}(\mathbf{r}, t) = \sum_{\mathbf{k}} E_{\mathbf{k}} \mathbf{e}_{\mathbf{k}}^{(\lambda)} \left[\alpha_{\mathbf{k}, \lambda} e^{i(\mathbf{k}\mathbf{r} - \omega_k t)} + \alpha_{\mathbf{k}, \lambda}^* e^{-i(\mathbf{k}\mathbf{r} - \omega_k t)} \right], \quad (1.6)$$

where \mathbf{k} is the index of the mode, λ the polarization, $\omega_{\mathbf{k}}$ the angular frequency of the mode \mathbf{k} , $\mathbf{e}_{\mathbf{k}}^{(\lambda)}$ the unit polarization vector, $\alpha_{\mathbf{k}}$ and $\alpha_{\mathbf{k}}^*$ are dimensionless complex constants and

$$E_{\mathbf{k}} = \left(\frac{\hbar \omega_{\mathbf{k}}}{2\epsilon_0} \right)^{1/2}, \quad (1.7)$$

contains all the dimensional prefactors. The magnetic field reads,

$$\mathbf{H}(\mathbf{r}, t) = \frac{1}{\mu_0} \sum_{\mathbf{k}} E_{\mathbf{k}} \frac{\mathbf{k} \times \mathbf{e}_{\mathbf{k}}^{(\lambda)}}{\omega_{\mathbf{k}}} \left[\alpha_{\mathbf{k}, \lambda} e^{i(\mathbf{k}\mathbf{r} - \omega_k t)} + \alpha_{\mathbf{k}, \lambda}^* e^{-i(\mathbf{k}\mathbf{r} - \omega_k t)} \right], \quad (1.8)$$

with the electric and magnetic field in phase and with directions orthogonal to one another and to the propagation direction.

Quantization

The radiation field is quantized by identifying $\alpha_{\mathbf{k},\lambda}$ and $\alpha_{\mathbf{k},\lambda}^*$ with the harmonic oscillator annihilation $\hat{a}_{\mathbf{k},\lambda}$ and creation $\hat{a}_{\mathbf{k},\lambda}^\dagger$ operators, which satisfy the commutation relation of bosons,

$$\begin{aligned} [\hat{a}_{\mathbf{k},\lambda}, \hat{a}_{\mathbf{k}',\lambda'}^\dagger] &= \delta_{\mathbf{k}\mathbf{k}'}\delta_{\lambda\lambda'}, \\ [\hat{a}_{\mathbf{k},\lambda}, \hat{a}_{\mathbf{k}',\lambda'}] &= 0, \\ [\hat{a}_{\mathbf{k},\lambda}^\dagger, \hat{a}_{\mathbf{k}',\lambda'}^\dagger] &= 0. \end{aligned} \quad (1.9)$$

The quantized electric and magnetic field take the form

$$\mathbf{E}(\mathbf{r}, t) = \sum_{\mathbf{k}} E_{\mathbf{k}} \mathbf{e}_{\mathbf{k}}^{(\lambda)} \left[\hat{a}_{\mathbf{k},\lambda} e^{i(\mathbf{k}\mathbf{r} - \omega_k t)} + \hat{a}_{\mathbf{k},\lambda}^\dagger e^{-i(\mathbf{k}\mathbf{r} - \omega_k t)} \right], \quad (1.10)$$

$$\mathbf{H}(\mathbf{r}, t) = \frac{1}{\mu_0} \sum_{\mathbf{k}} E_{\mathbf{k}} \frac{\mathbf{k} \times \mathbf{e}_{\mathbf{k}}^{(\lambda)}}{\omega_{\mathbf{k}}} \left[\hat{a}_{\mathbf{k},\lambda} e^{i(\mathbf{k}\mathbf{r} - \omega_k t)} + \hat{a}_{\mathbf{k},\lambda}^\dagger e^{-i(\mathbf{k}\mathbf{r} - \omega_k t)} \right]. \quad (1.11)$$

Fock States Representation

Single mode For a single mode of the field of frequency ω we define the creation and annihilation operators \hat{a}^\dagger and \hat{a} , respectively. The eigenstates $|n\rangle$ of the number operator $\hat{N} = \hat{a}^\dagger \hat{a}$ with eigenvalue n ,

$$\hat{a}^\dagger \hat{a} |n\rangle = n |n\rangle, \quad (1.12)$$

are called Fock states or photon number states and are usually interpreted as corresponding to the presence of n quanta of light in the corresponding mode. The states $|n\rangle$ are also eigenvectors of the Hamiltonian

$$H |n\rangle = \hbar\omega \left(\hat{a}^\dagger \hat{a} + \frac{1}{2} \right) |n\rangle = E_n |n\rangle, \quad (1.13)$$

with energy eigenvalue $E_n = \hbar\omega(n + 1/2)$.

Photon States The state containing no photons ($|0\rangle$) is called the vacuum state. Using the commutation relations (1.9) and the definition of the number operator ($\hat{N} = \hat{a}^\dagger \hat{a}$) one can derive the relations

$$\hat{a} |n\rangle = \sqrt{n} |n-1\rangle, \quad (1.14)$$

$$\hat{a}^\dagger |n\rangle = \sqrt{n+1} |n+1\rangle, \quad (1.15)$$

which, applying \hat{a}^\dagger successively on $|0\rangle$, give

$$|n\rangle = \frac{1}{\sqrt{n!}} (\hat{a}^\dagger)^n |0\rangle. \quad (1.16)$$

The Fock states ($|n\rangle$) being eigenstates of the number operator, they obviously form a complete basis of orthogonal states,

$$\langle n | m \rangle = \delta_{n,m} \text{ (Orthogonality)}, \quad (1.17)$$

$$\sum_n |n\rangle \langle n| = \mathbb{I} \text{ (Completeness relation)}. \quad (1.18)$$

In general an arbitrary superposition of energy eigenstates reads,

$$|\psi\rangle = \sum_n c_n |n\rangle. \quad (1.19)$$

More generally any state of one mode of light can be described by the density operator,

$$\rho = \sum_{n,m=0}^{\infty} \rho_{n,m} |n\rangle\langle m|, \quad (1.20)$$

where $\text{Tr}[\rho] = 1$ and ρ is a Hermitian (real eigenvalues λ_i) positive operator ($\lambda_i \geq 0$).

Multi-modes So far we have considered a single-mode field and have found that, the photon number states $\{|n\rangle\}$ form a basis of the Hilbert space. This formalism can be extended to multi-mode fields by defining the basis $|n_{\mathbf{k}}\rangle$,

$$|n_{\mathbf{k}}\rangle = |n_1\rangle \otimes |n_2\rangle \otimes \dots \otimes |n_k\rangle. \quad (1.21)$$

with n_1 photons in the first mode, n_2 in the second, and so forth.

Quadratures Operators

The definition of the electric field using the annihilation and creation operators (\hat{a}, \hat{a}^\dagger) given in (1.10) can be rewritten for a single mode as

$$\mathbf{E}(\mathbf{r}, t) = E_0 \mathbf{e} \left[\hat{x} \cos(\mathbf{k}\mathbf{r} - \omega_k t) + \hat{p} \sin(\mathbf{k}\mathbf{r} - \omega_k t) \right] \quad (1.22)$$

where the dimensionless operators \hat{x} and \hat{p} are the so-called quadratures of the electromagnetic field,

$$\hat{x} = \frac{1}{\sqrt{2}}(\hat{a}^\dagger + \hat{a}) \quad (1.23)$$

$$\hat{p} = \frac{i}{\sqrt{2}}(\hat{a}^\dagger - \hat{a}), \quad (1.24)$$

formally equivalent to the position and momentum of an harmonic oscillator. The operators \hat{x} and \hat{p} being Hermitian they can be measured as opposed to $(\hat{a}, \hat{a}^\dagger)$. They satisfy the commutation relation

$$[\hat{x}, \hat{p}] = i, \quad (1.25)$$

which gives the well-known Heisenberg uncertainty relation

$$\Delta \hat{x} \Delta \hat{p} \geq \frac{1}{2} |\langle [\hat{x}, \hat{p}] \rangle| = \frac{1}{2}. \quad (1.26)$$

with $\Delta A = (\langle A^2 \rangle - \langle A \rangle^2)^{1/2}$. The operators \hat{x} and \hat{p} being formally equivalent to the position and momentum, one can define, by analogy with classical mechanics, a phase-space representation of the quantum state of light.

Quadrature Eigenstates

The eigenstates of the quadratures,

$$\hat{x}|x\rangle = x|x\rangle, \quad (1.27)$$

$$\hat{p}|p\rangle = p|p\rangle, \quad (1.28)$$

form two sets of orthonormal states,

$$\langle x|x'\rangle = \delta(x-x'), \quad (1.29)$$

$$\langle p|p'\rangle = \delta(p-p'), \quad (1.30)$$

such as the position and momentum eigenstates. Both ensembles are a resolution of the identity,

$$\int_{-\infty}^{+\infty} |x\rangle\langle x| = \mathbb{I}, \quad (1.31)$$

$$\int_{-\infty}^{+\infty} |p\rangle\langle p| = \mathbb{I}, \quad (1.32)$$

which shows that they are both complete orthogonal bases. Both bases are related by a Fourier transformation,

$$|p\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dp e^{ixp} |x\rangle, \quad (1.33)$$

$$|x\rangle = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} dx e^{-ixp} |p\rangle. \quad (1.34)$$

The wave function and its Fourier transform of a given quantum state ψ read,

$$\psi(x) = \langle x|\psi\rangle, \quad (1.35)$$

$$\psi(p) = \langle p|\psi\rangle. \quad (1.36)$$

Coordinate Representation of Fock States

The coordinate representation of $|n\rangle$ is given by

$$\phi_n(x) = \langle x|n\rangle. \quad (1.37)$$

It follows from the definition (1.24) that

$$\hat{a}|0\rangle = \frac{1}{\sqrt{2}}(\hat{x} + i\hat{p})|0\rangle = \frac{1}{\sqrt{2}}\left(x + \frac{\partial}{\partial x}\right)\phi_0(x) = 0. \quad (1.38)$$

After solving the differential equation we obtain

$$\phi_0(x) = \frac{1}{\pi^{1/4}} e^{-x^2/2}. \quad (1.39)$$

The probability distribution of the x quadrature is a Gaussian of variance $\sigma^2 = 1/2$,

$$|\phi_0(x)|^2 = \frac{1}{\pi^{1/2}} e^{-x^2} = \frac{1}{(2\pi\sigma^2)^{1/2}} e^{-x^2/2\sigma^2}. \quad (1.40)$$

For higher photon numbers n we obtain the eigenstates of the harmonic oscillator

$$\phi_n(x) = \frac{1}{\sqrt{2^n n!}} H_n(x) \phi_0(x), \quad (1.41)$$

where $H_n(x)$ are the Hermite polynomials [2].

Moments of \hat{x} (same for \hat{p}) The mean value of an operator A on a quantum state of light ρ reads,

$$\langle A \rangle = \text{Tr}(\rho A). \quad (1.42)$$

In the particular case of Fock states the mean value of \hat{x} is null,

$$\langle \hat{x} \rangle = \langle n | \hat{x} | n \rangle \propto \langle n | \hat{a}^\dagger + \hat{a} | n \rangle = \langle n | n + 1 \rangle + \langle n | n - 1 \rangle = 0. \quad (1.43)$$

The second moment can be calculated in a similar way,

$$\langle \hat{x}^2 \rangle = \langle n | \hat{x}^2 | n \rangle = \frac{1}{2} \langle n | (\hat{a}^\dagger + \hat{a})^2 | n \rangle = \frac{1}{2} \langle n | (\hat{a}^{\dagger 2} + [\hat{a}, \hat{a}^\dagger] + 2\hat{a}^\dagger \hat{a} + \hat{a}^2) | n \rangle, \quad (1.44)$$

which gives, using (1.15),

$$\langle \hat{x}^2 \rangle = \frac{1}{2} \left(\langle n | n + 2 \rangle + 1 + 2n \langle n - 1 | n - 1 \rangle + \langle n | n - 2 \rangle \right) = n + \frac{1}{2}. \quad (1.45)$$

Notice that the same results are obtained for the p quadrature, which gives the uncertainty product,

$$\Delta \hat{x} \Delta \hat{p} = n + \frac{1}{2}, \quad (1.46)$$

which is minimum only for the vacuum state (saturating equation (1.26)).

1.2 Coherent States and Displacements

The number states form a useful representation of states with a small number of quanta. Unfortunately perfect number states are extremely difficult to generate experimentally for $n > 2$. On the other hand, lasers are very common sources of light, which generate the so-called coherent states.

Definitions A coherent state, denoted $|\alpha\rangle$, is an eigenstate of the annihilation operator,

$$\hat{a}|\alpha\rangle = \alpha|\alpha\rangle, \quad (1.47)$$

where α is a complex number (note that \hat{a} is a non-Hermitian operator). Alternatively, if one solves the Schrödinger equation for the light field emitted by a monochromatic dipole whose current oscillation is of frequency ω , one gets a coherent state [172]. We observe that a coherent state can be seen as a displaced vacuum state, where $D(\alpha)$ is the displacement operator,

$$|\alpha\rangle = e^{\alpha \hat{a}^\dagger - \alpha^* \hat{a}} |0\rangle = D(\alpha) |0\rangle. \quad (1.48)$$

The displacement operator being a unitary operator we obtain

$$D^\dagger(\alpha) = D^{-1}(\alpha) = D(-\alpha). \quad (1.49)$$

Properties Using the Baker-Campbell-Hausdorff formula provided that $[A, [A, B]] = 0$ and $[B, [A, B]] = 0$ we have

$$e^{A+B} = e^A e^B e^{-[A,B]/2}. \quad (1.50)$$

We can then rewrite the displacement operator in the normal and antinormal forms,

$$D(\alpha) = e^{-|\alpha|^2/2} e^{\alpha \hat{a}^\dagger} e^{-\alpha^* \hat{a}} \text{ (normal form)}, \quad (1.51)$$

$$D(\alpha) = e^{|\alpha|^2/2} e^{-\alpha^* \hat{a}} e^{\alpha \hat{a}^\dagger} \text{ (antinormal form)}. \quad (1.52)$$

Using the normal and antinormal forms and the Baker-Campbell-Hausdorff lemma

$$e^{-\alpha A} B e^{\alpha A} = B - \alpha[A, B] + \frac{\alpha^2}{2!}[A, [A, B]] + \dots \quad (1.53)$$

we obtain the action of the displacement operator on the annihilation and creation operators,

$$D^\dagger(\alpha) \hat{a} D(\alpha) = \hat{a} + \alpha \quad (1.54)$$

$$D^\dagger(\alpha) \hat{a}^\dagger D(\alpha) = \hat{a}^\dagger + \alpha^*, \quad (1.55)$$

where we see that $D(\alpha)$ displaces the operators \hat{a} (\hat{a}^\dagger) by an amount α (α^*). It is easy to prove that the definition of the coherent state as a displacement of the vacuum is equivalent to the definition as an eigenstate of \hat{a} :

$$\hat{a}|\alpha\rangle = D(\alpha) \underbrace{D^\dagger(\alpha) \hat{a} D(\alpha)}_{\hat{a} + \alpha} |0\rangle = \alpha D(\alpha) |0\rangle = \alpha |\alpha\rangle \quad (1.56)$$

as $\hat{a}|0\rangle = 0$.

Similarly (combining Eq. (1.55)) one can find the action of the displacement operator on the quadratures of the field,

$$D^\dagger(\alpha) \hat{x} D(\alpha) = \hat{x} + \sqrt{2} \Re \alpha \quad (1.57)$$

$$D^\dagger(\alpha) \hat{p} D(\alpha) = \hat{p} + \sqrt{2} \Im \alpha, \quad (1.58)$$

where \Re (\Im) is the real (imaginary) part.. We see that a coherent state ($|\alpha\rangle$) results from the displacement of the vacuum state ($|0\rangle$) in the phase space by an amount $d_x = \sqrt{2} \Re \alpha$ along the quadrature \hat{x} and $d_p = \sqrt{2} \Im \alpha$ along \hat{p} .

Expansion in terms of number states Using the normal form of the displacement operator we obtain

$$|\alpha\rangle = e^{-|\alpha|^2/2} e^{\alpha \hat{a}^\dagger} e^{-\alpha^* \hat{a}} |0\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n (\hat{a}^\dagger)^n}{n!} |0\rangle, \quad (1.59)$$

which finally gives,

$$|\alpha\rangle = e^{-|\alpha|^2/2} \sum_{n=0}^{\infty} \frac{\alpha^n}{\sqrt{n!}} |n\rangle. \quad (1.60)$$

We see that coherent states have an undefined number of photons, which allows them to have a better defined phase than number states (which have totally random phase). The probability of finding n photons in $|\alpha\rangle$ is given by a Poisson distribution,

$$p(n) = |\langle n | \alpha \rangle|^2 = e^{-|\alpha|^2} \frac{|\alpha|^{2n}}{n!}, \quad (1.61)$$

with mean value and variance such that $\langle n \rangle = \Delta^2 n = |\alpha|^2$.

Quasi-Classical States Coherent states are often called quasi-classical states as the product of uncertainties in amplitude and phase is the minimum allowed by the uncertainty principle (1.26), with

$$\Delta^2 \hat{x} = \Delta^2 \hat{p} = \frac{1}{2}, \quad (1.62)$$

being the closest to a classical state where one knows exactly both the amplitude and phase. The proof is similar to the calculation of the second order moment in (1.1),

$$\langle \hat{x}^2 \rangle = \langle \alpha | (\hat{a}^{\dagger 2} + [\hat{a}, \hat{a}^\dagger] + 2\hat{a}^\dagger \hat{a} + \hat{a}^2) | \alpha \rangle = \frac{1}{2} (1 + (\alpha + \alpha^*)^2) = \frac{1}{2} + \langle \hat{x} \rangle^2, \quad (1.63)$$

which combined with $\Delta^2 \hat{x} = \langle \hat{x}^2 \rangle - \langle \hat{x} \rangle^2$ gives the desired result.

Overcomplete Basis Using the definition (1.48) and property (1.49) we obtain,

$$\langle \beta | \alpha \rangle = \langle 0 | D^\dagger(\beta) D(\alpha) | 0 \rangle = \langle 0 | D(\alpha - \beta) | 0 \rangle = \langle 0 | \alpha - \beta \rangle = e^{-|\alpha - \beta|^2/2}, \quad (1.64)$$

giving finally,

$$|\langle \beta | \alpha \rangle|^2 = e^{-|\alpha - \beta|^2}. \quad (1.65)$$

Thus coherent states are not orthogonal, but become approximately orthogonal in the limit where α is far from β . The set of coherent states satisfy the completeness relation,

$$\frac{1}{\pi} \int |\alpha\rangle \langle \alpha| d^2\alpha = \mathbb{I}. \quad (1.66)$$

This can be proved by using the expansion of coherent states in terms of number states (1.60), the polar coordinates $\alpha = re^{i\theta}$ and the integrals,

$$\int_0^{2\pi} e^{i(n-m)\theta} d\theta = 2\pi \delta_{n,m}, \quad (1.67)$$

and

$$\int_0^\infty dx e^{-x} x^n = n!, \quad (1.68)$$

as shown in [172].

1.3 Linear Optical Operations

In the following we are going to detail the ensemble of linear operations that can be applied to a multimode optical field. In most of the cases we will introduce the Hamiltonian of the interaction using the creation and annihilation operators, as their derivation is more intuitive. In some cases we will derive the transformation just for the quadratures (\hat{x}, \hat{p}) , as they are physical quantities that can be measured in contrast to creation and annihilation operators.

Heisenberg Equation of Motion

For an operator A which does not depends explicitly on time, the Heisenberg's equation of motion reads

$$\frac{dA}{dt} = \frac{1}{i\hbar} [A, H], \quad (1.69)$$

where H is the Hamiltonian describing the system in the Heisenberg picture.

Phase Shift

The phase is the crucial element of the wave behavior of the electromagnetic field. The phase of a single beam has no physical meaning, as it cannot be measured since we have only access to the phase difference between different beams. A usual way of applying a phase shift on a optical beam is to increase the path length of the beam compared to the others. Adding some transparent material with refractive index higher than vacuum on the path of the beam has a similar effect.

In both situations the beam accumulates an additional phase θ proportional to the path difference or the interaction time ($\theta = \omega\Delta t$) respectively. The corresponding Hamiltonian (H_θ) is simply the Hamiltonian of the free electromagnetic field,

$$H_\theta = \hbar\omega \hat{a}^\dagger \hat{a}. \quad (1.70)$$

Using the Heisenberg equation of motion (1.69) we obtain

$$\frac{d\hat{a}}{dt} = -i\omega[\hat{a}, \hat{a}^\dagger \hat{a}] = i\omega\hat{a}, \quad (1.71)$$

which is a first order differential equation, that combined with the initial condition $\hat{a}(0) = \hat{a}_{in}$ gives

$$\hat{a}_{out} = e^{-i\omega\Delta t} \hat{a}_{in} = e^{-i\theta} \hat{a}_{in}, \quad (1.72)$$

and similarly, $\hat{a}_{out}^\dagger = e^{i\theta} \hat{a}_{in}^\dagger$. Using the same technique with $H_\theta = \frac{\hbar\omega}{2}[\hat{x}^2 + \hat{p}^2 - 1]$ and using the initial conditions ($\hat{x}(0) = \hat{x}_{in}, \hat{p}(0) = \hat{p}_{in}$) one obtains the quadratures transformation,

$$\begin{bmatrix} \hat{x} \\ \hat{p} \end{bmatrix}_{out} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{p} \end{bmatrix}_{in}, \quad (1.73)$$

which is just a rotation of angle θ in the phase-space representation.

Beamsplitter

A beamsplitter is a semi-transparent mirror which transmits part of the incoming signal and the rest is reflected. As shown in Fig. 1.1 when two beams are spatially and temporally matched in a beamsplitter the outgoing modes are a coherent mixture of both input modes. The interaction between two beams on a beamsplitter has the

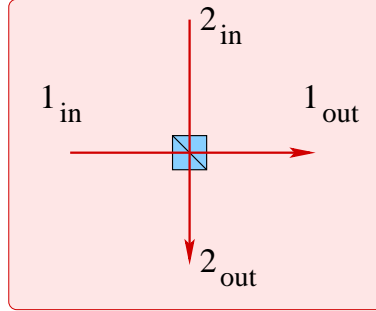


Figure 1.1: Two light modes (1 and 2) are matched at the input ports of a beamsplitter which outputs two linear combination of the input modes.

same effect as switching on the interaction described by the following Hamiltonian

$$H_\gamma = \hbar\omega(\hat{a}_2^\dagger \hat{a}_1 + \hat{a}_1^\dagger \hat{a}_2), \quad (1.74)$$

during a given time Δt , which coherently mixes both modes while preserving the total number of photons ($\hat{N}_1 + \hat{N}_2$). Using the Heisenberg equation of motion (1.69) we obtain the system of differential equations of the annihilation operators of both modes (\hat{a}_1, \hat{a}_2),

$$\frac{d\hat{a}_1}{dt} = -i\omega[\hat{a}_1, \hat{a}_1^\dagger \hat{a}_2] = -i\omega\hat{a}_2, \quad (1.75)$$

$$\frac{d\hat{a}_2}{dt} = -i\omega[\hat{a}_2, \hat{a}_2^\dagger \hat{a}_1] = -i\omega\hat{a}_1, \quad (1.76)$$

which reduces to the second order differential equation for \hat{a}_1

$$\frac{d^2\hat{a}_1}{dt^2} = -\omega^2\hat{a}_1, \quad (1.77)$$

with a similar equation for \hat{a}_2 . The solution of the differential equation gives us the transformation,

$$\begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \end{bmatrix}_{out} = \begin{bmatrix} \cos(\omega\Delta t) & -i\sin(\omega\Delta t) \\ -i\sin(\omega\Delta t) & \cos(\omega\Delta t) \end{bmatrix} \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \end{bmatrix}_{in}. \quad (1.78)$$

In the case of the beamsplitter the term $\cos(\omega\Delta t)$ ($\sin(\omega\Delta t)$) becomes the square root of the transmittance (reflectance) of the beamsplitter,

$$\begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \end{bmatrix}_{out} = \begin{bmatrix} \sqrt{T} & -i\sqrt{R} \\ -i\sqrt{R} & \sqrt{T} \end{bmatrix} \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \end{bmatrix}_{in}. \quad (1.79)$$

The phase i on the non-diagonal terms results from the boundary condition on a semi-transparent mirror which implies that reflected waves get a phase i with respect to transmitted waves. In the quantum optics literature the beamsplitter transformation is currently written

$$\begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \end{bmatrix}_{out} = \begin{bmatrix} \sqrt{T} & \sqrt{R} \\ \sqrt{R} & -\sqrt{T} \end{bmatrix} \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \end{bmatrix}_{in}, \quad (1.80)$$

which can be derived from the previous equation by applying the change of variable $\hat{a}_2 \rightarrow i\hat{a}_2$. The minus sign can be chosen arbitrarily in front of the \sqrt{T} term of the first or second line, both being equivalent up to a local phase.

The transformation of the quadratures in vector notation $\hat{r} = (\hat{x}, \hat{p})^T$ can be calculated in a similar way and reads,

$$\begin{bmatrix} \hat{r}_1 \\ \hat{r}_2 \end{bmatrix}_{out} = \begin{bmatrix} \sqrt{T}\mathbb{I} & \sqrt{R}\mathbb{I} \\ -\sqrt{R}\mathbb{I} & \sqrt{T}\mathbb{I} \end{bmatrix} \begin{bmatrix} \hat{r}_1 \\ \hat{r}_2 \end{bmatrix}_{in}, \quad (1.81)$$

where \mathbb{I} is here the identity in \mathbb{C}^2 . The minus sign can be chosen arbitrarily in front of the \sqrt{R} term of the first or second line, both being equivalent up to a local phase.

Displacement

Previously we have defined the displacement operator $D(\alpha)$ without suggesting any physical way of realizing it. The usual way of implementing the displacement operator in the lab is shown in Fig. 1.2. The idea is to combine the optical mode we want to

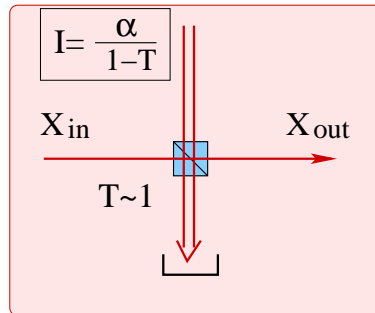


Figure 1.2: In order to apply a displacement α to a given optical mode. We combine the target mode \hat{x}_{in} with an auxiliary mode of amplitude $\alpha/\sqrt{1-T}$ into a beamsplitter with high transmittance.

displace (quadrature \hat{x}_{in}) with an auxiliary mode consisting on a high intensity coherent

state of mean value $\bar{d}/\sqrt{1-T}$, where $\bar{d} = \sqrt{2}(\Re\alpha, \Im\alpha) = (d_x, d_p)$ into a beamsplitter of high transmittance ($T \rightarrow 1$).

Using the transformation of the beamsplitter for the quadratures of the field (1.81) the output quadrature \hat{x}_{out} reads (similarly for \hat{p}_{out}),

$$\hat{x}_{out} = \sqrt{T}\hat{x}_{in} + \sqrt{1-T}\hat{x}_{aux} = \sqrt{T}\hat{x}_{in} + d_x. \quad (1.82)$$

On the other hand the variance of the output mode reads,

$$\Delta^2\hat{x}_{out} = T\Delta^2\hat{x}_{in} + \frac{(1-T)}{2}, \quad (1.83)$$

as the auxiliary beam being a coherent state we have $\Delta^2\hat{x}_{aux} = 1/2$. Equation (1.82) shows that in the limit of very high transmittance ($T \rightarrow 1$) the output quadrature is exactly the input quadrature displaced by \bar{d} . Looking at equation (1.83) we observe that the higher the transmittance is, the less the auxiliary beam disturbs the state. The price to pay for a highly efficient displacement is the increase in the intensity of the auxiliary beam, as $\alpha/\sqrt{1-T} \rightarrow \infty$ when $T \rightarrow 1$. There is a clear tradeoff between the efficiency of the displacement operation on one side and the intensity of the auxiliary beam we can reach and the transmittance of the beamsplitter on the other side.

Measurement

In quantum mechanics one can associate a measurement to each basis that is a resolution of the identity. In quantum optics we consider two types of measurement, those that resolve the photon number states and those who measure the quadratures of the field, as we present below.

Single Photon Sensitive Detectors

The measurement related to the Fock basis (photon number states) is the so-called detector with photon resolution as it is capable of discriminating among all Fock states ($|n\rangle\langle n|$). Unfortunately discriminating the photon number is so extremely challenging that there is no actual detector capable of doing it efficiently. A more reasonable task is the so-called detector with photon sensitivity that is capable of resolving between either no photon ($|0\rangle\langle 0|$) or one or more photons ($\mathbb{I} - |0\rangle\langle 0|$). Photon sensitivity is currently achieved using avalanche photodiodes (APD) which are tuned to sense a single photon, which is already technically very challenging.

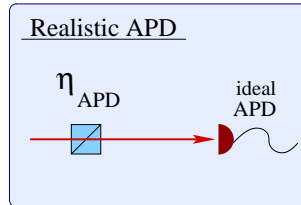


Figure 1.3: A realistic APD with efficiency η_{APD} is modeled by placing a beamsplitter of transmittance η_{APD} before an ideal APD detector.

Realistic APD A real APD has two sources of error. Firstly not all the photons arriving at the detector generate an avalanche. The rate of detected compared to arriving photons is called the efficiency (η_{APD}) of the detector. It is modeled by a

beamsplitter of transmittance η_{APD} placed before a perfect detector, as shown in Fig. 1.3. Current APD detector technology reaches around 50% of efficiency at most. The second source of errors is the so-called dark counts, as they correspond to spontaneous clicks not heralded by any impinging photon. Fortunately, the effect of dark counts can be reduced to a negligible value if the detector is triggered only when a pulse is expected.

Homodyne Detection

The way of measuring the quadratures of the electromagnetic field is the so-called homodyne detection shown in Fig. 1.4. The target mode (\hat{x}_t, \hat{p}_t) is combined with a

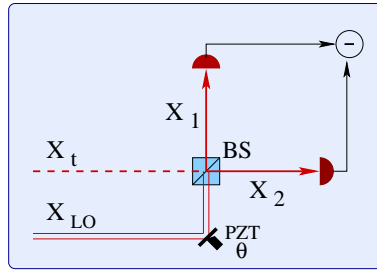


Figure 1.4: The target mode \hat{x}_t is combined with the local oscillator X_{LO} into a balanced beamsplitter. The intensity of the outgoing modes are measured with two photodetectors, which after subtraction give a signal proportional to the measured quadrature \hat{x}_t . In order to measure among another quadrature \hat{x}_θ we have to apply a phase shift θ using for example a piezoelectric transducer (PZT) to the local oscillator.

so-called local oscillator (LO) X_{LO} into a balanced beamsplitter. The local oscillator is the phase reference of the system, being a classical beam ($\sim 10^9$ photons), where we wrote X_{LO} for the quadrature of the local oscillator in order to stress its classical nature. The local oscillator being the phase reference we can fix without loss of generality the local oscillators quadratures to $(X_{LO}, 0)$, then the outgoing modes 1, 2 read

$$\hat{x}_1 = (\hat{x}_t + X_{LO})/\sqrt{2} \quad (1.84)$$

$$\hat{p}_1 = \hat{p}_t/\sqrt{2} \quad (1.85)$$

$$\hat{x}_2 = (\hat{x}_t - X_{LO})/\sqrt{2} \quad (1.86)$$

$$\hat{p}_2 = \hat{p}_t/\sqrt{2}. \quad (1.87)$$

The intensities of the outgoing modes are then measured using two photodiodes,

$$I_{1,2} = k\hat{N}_{1,2} = \frac{k}{2}(\hat{x}_{1,2}^2 + \hat{p}_{1,2}^2 - 1), \quad (1.88)$$

where the constant k contains all the dimensional prefactors. The two photocurrents $I_{1,2}$ are subsequently subtracted and amplified with a low noise amplifier in order to obtain an estimation of the quadrature \hat{x}_t ,

$$I_1 - I_2 = \frac{k}{2}((\hat{x}_t + X_{LO})^2 - (\hat{x}_t - X_{LO})^2) = kX_{LO}\hat{x}_t. \quad (1.89)$$

The local oscillator being classical its intensity kX_{LO}^2 can be estimated without disturbing it, allowing one to calculate \hat{x}_t from the difference of the photocurrents and the intensity of the local oscillator. In order to measure the conjugate quadrature \hat{p}_t

we apply a phase shift of $\pi/2$ to the local oscillator transforming the local oscillator to $(0, P_{LO})$ which after subtraction of the photocurrents gives the quadrature \hat{p}_t . In full generality one can homodyne any quadrature $\hat{x}_\theta = \cos\theta\hat{x} + \sin\theta\hat{p}$ by applying a phase shift θ to the local oscillator, using for example a piezoelectric transducer (PZT) which changes the path length of the local oscillator compared to the target mode.

The fact that the beam impinging on the photodiodes is classical, due to the intensity of the local oscillator, strikingly simplifies the setup as we only need to use pin photodiodes. In order to successfully implement a quantum homodyne measurement the noise added by the electronics (amplifier and subtraction step) must be far below the shot noise in order to be able to distinguish the quantum noise. As an example, the different homodyne detections implemented by the group of P. Grangier in Orsay reach an electronic noise which is 20dB below the shot noise [94, 127]. Despite being technically challenging, homodyne detection can reach extremely high detection efficiencies of 90% [94, 152, 203].

Realistic homodyne detection The efficiency (η_{BHD}) of an homodyne detection is modeled by placing a beamsplitter of transmittance η_{BHD} before an ideal homodyne detection, as shown in Fig. 1.5. The quadrature (\hat{x}_m) of the impinging mode reads,

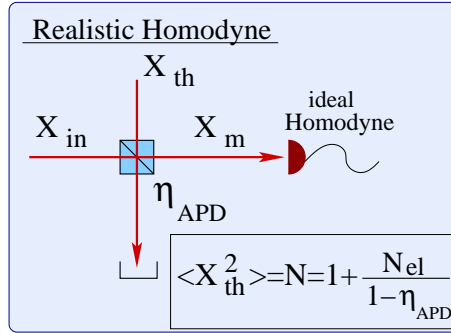


Figure 1.5: A realistic homodyne detection with efficiency η_{BHD} and electronic noise N_{el} is modeled by placing a beamsplitter of transmittance η_{BHD} and a thermal state of variance $N_{el}/(1 - \eta_{BHD}) + 1$ added at the second port of the beamsplitter before an ideal homodyne detector. Notice that in order to simplify the scheme we have represented the ideal homodyning using a single detector (omitting the local oscillator) instead of reproducing the scheme of Fig. 1.4, as done in most of theoretical works.

$$\hat{x}_m = \sqrt{\eta_{BHD}}\hat{x}_{in} + \sqrt{1 - \eta_{BHD}}\hat{x}_{th}. \quad (1.90)$$

The attenuation can be compensated by applying a rescaling of factor $1/\sqrt{\eta_{BHD}}$ to the measured quadrature \hat{x}_m , the added noise referred to the input then reads,

$$\chi_D = \frac{1 - \eta_{BHD}}{\eta_{BHD}}. \quad (1.91)$$

In addition to the imperfect detection efficiency η_{BHD} , the electronic noise of the homodyne detector is another factor that may reduce the quality of the measurement. We model the added noise by assuming that the effective quadrature \hat{x}_m is combined in the beamsplitter η_{BHD} with a thermal state \hat{x}_{th} of variance $N/2$,

$$\hat{x}_m = \sqrt{\eta_{BHD}}\hat{x}_{in} + \sqrt{1 - \eta_{BHD}}\hat{x}_{th}. \quad (1.92)$$

The electronic noise N_{el} corresponds to thermal photons that arrive to the ideal detector. It reads,

$$N_{el} = (1 - \eta_{BHD})(N - 1), \quad (1.93)$$

where the added noise referred to the input reads,

$$\chi_D = \frac{1 + N_{el}}{\eta_{BHD}} - 1. \quad (1.94)$$

1.4 Non-linear Optical Operations

The invention of the laser, delivering high intensity monochromatic light made possible the observation of nonlinear optical processes. In a nonlinear media the dielectric polarization vector is written as a power series in the electrical field

$$P(t) \propto \chi^{(1)}E(t) + \chi^{(2)}E^2(t) + \chi^{(3)}E^3(t) + \dots \quad (1.95)$$

where the $\chi^{(n)}$ are the n -th order susceptibilities of the medium. The linear optics transformations presented above correspond to the $\chi^{(1)}$ term, while the high order terms yields non-linear effects.

Second order nonlinear effects $\chi^{(2)}$

When the incident field enters a nonlinear medium the second order nonlinear component of the polarization can generate four different effects (see in Fig. 1.6):

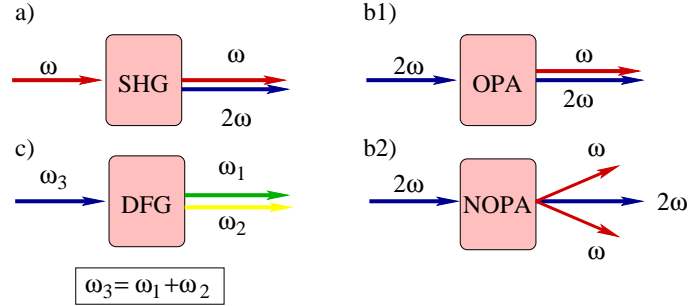


Figure 1.6: Second order nonlinear processes: a) Second Harmonic Generation (SHG). b) Optical Parametric Amplification (OPA): degenerate (b1) and non-degenerate (NOPA) (b2). c) Difference Frequency Generation (DFG).

1. Second Harmonic Generation (SHG) is a process in which pairs of incident photons of energy $\hbar\omega$ ("red photons") interacting with the nonlinear material are effectively combined to form new photons with the energy $2\hbar\omega$ ("blue photons").
2. Optical Parametric Amplification (OPA) can be seen as the time reversal of SHG, where one photon of energy $2\hbar\omega$ is splitted into two photons of energy $\hbar\omega$. The OPA can be degenerated if the two outgoing photons are generated on the same mode, and non degenerated (NOPA) if the two photons are generated on independent modes.
3. Difference Frequency Generation (DFG) is a generalization of OPA generating photons with different energies $\hbar\omega_1$ and $\hbar\omega_2$ as shown in Fig. 1.6.
4. Sum Frequency Generation (SFG) is the symmetric counterpart of DFG.

Practical $\chi^{(2)}$ effects are obtained placing a special transparent crystal without inversion symmetry in a laser beam under suitable angle. The no inversion symmetry is crucial as inversion symmetry gives no second order term in the electric field. The phase matching conditions determine which of the four different second order effects listed previously will be generated. Phase matching conditions can be obtained by correctly orienting a highly birefringent crystal. Other techniques such as temperature tuning or quasi-phase-matching using periodically-poled crystals are also currently used.

Third order nonlinear effects $\chi^{(3)}$

The third order nonlinear processes are even weaker than the second order ones, but can be observed in material with inversion symmetry, as in this situation the second order is null. High intensity beams are necessary in order to observe $\chi^{(3)}$ effects such as the Kerr effect, self-phase modulation and optical solitons.

Optical Parametric Amplification

In the following we will concentrate our attention to $\chi^{(2)}$ nonlinear effects, more precisely to optical parametric amplification which allows one to generate squeezed vacuum states when working in a degenerate regime and two-mode squeezed vacuum when working in the non-degenerate regime, generating a rich family of states extremely useful in quantum information with continuous variables.

Squeezed states

When we pump a degenerate OPA with a bright laser, some of the pump photons of energy $2\hbar\omega$ are splitted into two $\hbar\omega$ photons. The OPA working in a degenerate regime, the outgoing mode must then be uniquely composed of even Fock states ($|2n\rangle$) or coherent superposition of such even photon numbers. The corresponding Hamiltonian must then contain a $\hat{a}^{\dagger 2}$ term in order to generate pairs of photons and a similar term \hat{a}^2 to ensure Hermiticity,

$$H = i\frac{\tau}{2}[e^{-i\phi}\hat{a}^2 - e^{i\phi}\hat{a}^{\dagger 2}], \quad (1.96)$$

where τ is the squeezing factor related to the intensity of the pump laser and the strength of the non-linear interaction and ϕ corresponds to a phase rotation. Using the Heisenberg equation of motion (1.69) we obtain the action of the squeezing operation on the annihilation and creation operators,

$$\begin{bmatrix} \hat{a} \\ \hat{a}^\dagger \end{bmatrix}_{out} = \begin{bmatrix} \cosh r & -e^{i\phi} \sinh r \\ -e^{-i\phi} \sinh r & \cosh r \end{bmatrix} \begin{bmatrix} \hat{a}_1 \\ \hat{a}_1^\dagger \end{bmatrix}_{in}, \quad (1.97)$$

where $r = \tau\Delta t$.

Quadratures The effect of the squeezing operation is better understood using the quadratures description of the electromagnetic field. The Hamiltonian (1.96), choosing $\phi = 0$ can also be written,

$$H = \frac{i}{2}[\hat{x}\hat{p} + \hat{p}\hat{x}], \quad (1.98)$$

which together with the Heisenberg equation of motion gives, after solving a simple differential equation, the quadratures transformation,

$$\begin{bmatrix} \hat{x} \\ \hat{p} \end{bmatrix}_{out} = \begin{bmatrix} e^{-r} & 0 \\ 0 & e^r \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{p} \end{bmatrix}_{in}, \quad (1.99)$$

where we clearly see that the effect of the operation is to squeeze one quadrature and anti-squeeze the conjugate quadrature. An arbitrary squeezing transformation $S(r, \phi)$ over a given quadrature \hat{x}_ϕ can be implemented by first applying a first phase shift ($P(-\phi)$) followed by a squeezing along \hat{x} and a final shift ($P(\phi)$) ($S(r, \phi) = P(\phi)S(r)P(-\phi)$).

Squeezed Vacuum When we pump a degenerate OPA with a bright laser of photons $2\hbar\omega$, the initial state of the field $\hbar\omega$ being just the vacuum, we obtain the so-called squeezed vacuum state $S(r)|0\rangle$. For squeezing factors $r > 0$ the variance of the squeezed quadrature decreases below the shot noise unit ($e^{-2r} < 1$). In order to satisfy the Heisenberg uncertainty relation the variance of the conjugate quadrature must increase with the squeezing ($e^{2r} > 1$), as shown in Fig. 1.7. In the Fock basis the squeezed

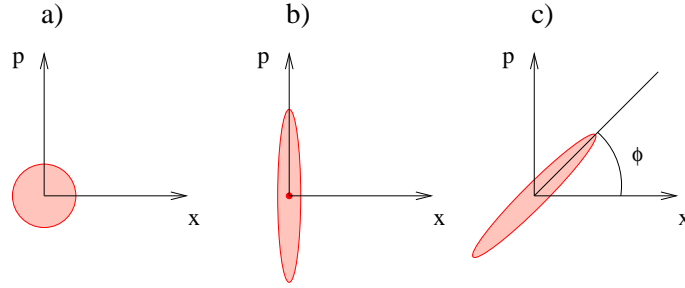


Figure 1.7: Optical Parametric Amplification of the vacuum. a) The initial vacuum state. b) Squeezed vacuum along the quadrature \hat{x} . c) Squeezed vacuum along the rotated quadrature \hat{x}_ϕ .

vacuum state reads,

$$S(r)|0\rangle = \frac{1}{\sqrt{\cosh r}} \sum_{n=0}^{\infty} \frac{\sqrt{(2n)!}}{2^n n!} \tanh r^n |2n\rangle, \quad (1.100)$$

where we observe that there is no odd Fock state, as expected. The mean number of photons ($\langle \hat{N} \rangle$) in the squeezed states can be easily calculated using the definition of the number operator $\hat{N} = (\hat{x}^2 + \hat{p}^2) - 1/2$ and equation (1.99) giving,

$$\langle \hat{N} \rangle = \frac{1}{2} [\langle (\hat{x}^2 + \hat{p}^2) \rangle - 1] = \sinh^2 r. \quad (1.101)$$

Squeezed Coherent State When we spatially and temporally match a bright laser of photons $2\hbar\omega$ with a coherent state of the field $\hbar\omega$ into a degenerate OPA we obtain the so-called squeezed coherent state $S(r)|\alpha\rangle$. As we shown in Fig. 1.8 the squeezing reduces the mean value of the squeezed quadrature and increases the conjugate one,

$$\langle \hat{x} \rangle_{out} = e^{-r} \langle \hat{x} \rangle_{in} \quad (1.102)$$

$$\langle \hat{p} \rangle_{out} = e^r \langle \hat{p} \rangle_{in}, \quad (1.103)$$

where the second moments are changed in the same way as for the squeezed vacuum state. Matching the pump beam and the signal coherent state into the OPA is a challenging task. An equivalent and simpler way of generating squeezed coherent states consists in applying a displacement operation ($D(\alpha)$) over a squeezed vacuum state ($D(\alpha)S(r)|0\rangle$).

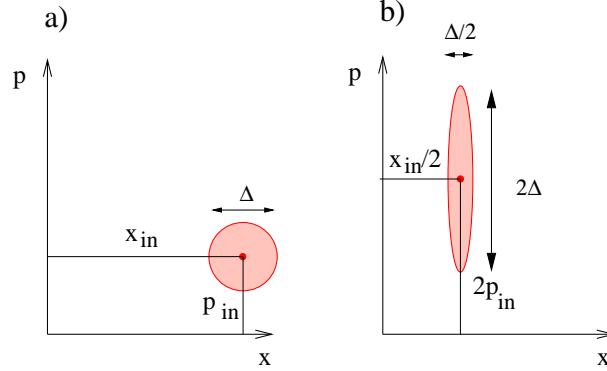


Figure 1.8: A squeezing operation with $e^r = 2$ applied to a coherent state displaces it to a new mean value $(e^{-r}\langle\hat{x}\rangle_{in}, e^r\langle\hat{p}\rangle_{in})$, squeezing the uncertainty $\Delta^2\hat{x}$ and stretching $\Delta^2\hat{p}$.

Two-modes Squeezed States

When we pump a non-degenerate OPA (NOPA) with a bright laser some of the pump photons of energy $2\hbar\omega$ are splitted into two $\hbar\omega$ photons which are emitted on different modes, usually called signal and idler. The number of photons on both modes must then be the same ($|n, n\rangle$), or a superposition of such states. The corresponding Hamiltonian must contain a term $\hat{a}_1^\dagger\hat{a}_2^\dagger$ in order to generate pairs of photons over different modes, reads,

$$H = i\tau[\hat{a}_1\hat{a}_2 - \hat{a}_1^\dagger\hat{a}_2^\dagger], \quad (1.104)$$

which can also be written using the quadratures of the fields,

$$H = i\tau[\hat{x}_1\hat{p}_2 + \hat{p}_1\hat{x}_2]. \quad (1.105)$$

Using the Heisenberg equation of motion (1.69), after solving a simple system of differential equations, we obtain the action of the two-mode squeezing operator $S_{TMS}(r)$ ($r = \tau\Delta t$) on the annihilation and creation operators of both fields,

$$\begin{bmatrix} \hat{x}_1 \\ \hat{p}_1 \\ \hat{x}_2 \\ \hat{p}_2 \end{bmatrix}_{out} = \begin{bmatrix} \cosh r & 0 & \sinh r & 0 \\ 0 & \cosh r & 0 & -\sinh r \\ \sinh r & 0 & \cosh r & 0 \\ 0 & -\sinh r & 0 & \cosh r \end{bmatrix} \begin{bmatrix} \hat{x}_1 \\ \hat{p}_1 \\ \hat{x}_2 \\ \hat{p}_2 \end{bmatrix}_{in}. \quad (1.106)$$

One can see looking at equation (1.106) that the superposition of the input quadratures $(\hat{x}_1 + \hat{x}_1)$ and $(\hat{p}_1 - \hat{p}_1)$ is anti-squeezed, where $(\hat{x}_1 - \hat{x}_1)$ and $(\hat{p}_1 + \hat{p}_1)$ it is squeezed.

Two mode squeezed vacuum When we pump a NOPA with a bright laser of photons $2\hbar\omega$, the signal and idler input fields being just the vacuum, we obtain the so-called two-mode squeezed vacuum state $S_{TMS}(r)|0, 0\rangle$. In the Fock basis the two-mode squeezed vacuum state reads,

$$S_{TMS}(r)|0, 0\rangle = \frac{1}{\sqrt{\cosh r}} \sum_{n=0}^{\infty} (\tanh r)^n |n, n\rangle. \quad (1.107)$$

In the limit of infinite squeezing we obtain perfect correlation among the \hat{x} quadratures ($\hat{x}_1 = \hat{x}_2$) and perfect anti-correlation for \hat{p} quadrature ($\hat{p}_1 = -\hat{p}_2$), which is nothing else than the well known EPR entangled state used in the seminal paper by Einstein,

Podolsky and Rosen (EPR) [69]. In the following we will use sometimes the term "EPR states" to mention all the two-mode squeezed states, not only those with infinite energy.

Amplification When we pump a NOPA with a bright laser of photons $2\hbar\omega$ and match it spatially and temporally with another signal ($\hbar\omega$), we obtain at the output an amplification of the $\hbar\omega$ signal. The amplification gain can be seen to be $G = \cosh^2 r$. The amplification is accompanied with an added noise $G - 1 = \sinh^2 r$. This optical amplification will be studied in more detail in the next chapter.

Thermal States

If we trace out one of the two output modes of a two-mode squeezed (1.107) state, we obtain the mixed state,

$$\rho = \text{Tr}|\psi\rangle\langle\psi|_{EPR} = \frac{1}{\cosh r} \sum_{n=0}^{\infty} (\tanh r)^{2n} |n\rangle\langle n|. \quad (1.108)$$

This state is just a thermal state with a Bose-Einstein distribution. From the estimation of the mean photon number,

$$\langle n \rangle = \langle \hat{N} \rangle = \cosh r - 1, \quad (1.109)$$

one can write $\tanh^2 r = \langle n \rangle / (\langle n \rangle + 1)$, which gives,

$$\rho_{TH} = \sum_{n=0}^{\infty} \frac{\langle n \rangle^n}{(\langle n \rangle + 1)^{n+1}} |n\rangle\langle n|, \quad (1.110)$$

which is exactly the equation of a thermal state.

Entanglement

The EPR state ($|\psi\rangle_{EPR}$) being a pure state and his partial traced state being a mixed state (thermal state (ρ_{TH})), it is a clear signature of the presence of entanglement in the EPR state. In this thesis we will study different properties of the EPR states as they are a key resource for quantum information processing and a source of non-local effects.

Chapter 2

Phase-Space Representation

2.1 Introduction to Continuous Variables

A continuous variable (CV) system is a canonical infinite dimensional quantum system composed of an ensemble of N modes described by a Hilbert space

$$\mathcal{H} = \bigotimes_{i=1}^N \mathcal{H}_i \quad (2.1)$$

resulting from the tensor product of N infinitely-dimensional Fock spaces \mathcal{H}_i . One could think of N modes of the electromagnetic field, where the modes can be distinguished either by having different energies (ω_i), polarizations or spatial modes. The space \mathcal{H}_i is spanned by the fock basis $\{|n\rangle_k\}$ of eigenstates of the number operator $\hat{n}_i = \hat{a}_i^\dagger \hat{a}_i$. The vacuum state of the global Hilbert space reads $|0\rangle = \bigotimes_i |0\rangle_i$, where $\hat{a}_i |0\rangle_i = 0$, is the ground state of the interaction-free Hamiltonian of a system of N harmonic oscillators,

$$H = \sum_{i=1}^N \left[\hat{a}_i^\dagger \hat{a}_i + \frac{1}{2} \right], \quad (2.2)$$

where \hat{a}_i and \hat{a}_i^\dagger are the annihilation and creation operators of mode i which satisfies the bosonic commutation relation,

$$[\hat{a}_i, \hat{a}_j^\dagger] = \delta_{ij}, \quad [\hat{a}_i, \hat{a}_j] = [\hat{a}_i^\dagger, \hat{a}_j^\dagger] = 0. \quad (2.3)$$

The corresponding quadrature operators for each mode are defined as

$$\hat{x} = \frac{1}{\sqrt{2}}(\hat{a}^\dagger + \hat{a}), \quad (2.4)$$

$$\hat{p} = \frac{i}{\sqrt{2}}(\hat{a}^\dagger - \hat{a}). \quad (2.5)$$

The quadratures can be grouped together in a vector \hat{r}

$$\hat{r} = (\hat{r}_1, \dots, \hat{r}_{2N})^T = (\hat{x}_1, \hat{p}_1, \hat{x}_2, \hat{p}_2, \dots, \hat{x}_N, \hat{p}_N)^T, \quad (2.6)$$

which enables us to write in a compact form the bosonic canonical commutation relations (CCR) between the quadratures operators,

$$[\hat{r}_k, \hat{r}_l] = i\Omega_{kl}, \quad (2.7)$$

where Ω is the symplectic form

$$\Omega = \bigoplus_{i=1}^N \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}. \quad (2.8)$$

Weyl Operators

In order to define the phase-space representation we need to introduce the so called Weyl operator,

$$D_\xi = e^{-i\xi^T \Omega \hat{r}}, \quad (2.9)$$

which is nothing else than the generalization to N modes of the displacement operator

$$D(\alpha) = e^{\alpha \hat{a}_i^\dagger - \alpha^* \hat{a}_i}, \quad (2.10)$$

redefined using the quadratures representation $((\xi_x, \xi_p) = \sqrt{2}(\Re \alpha, \Im \alpha))$,

$$D(\xi) = e^{i(d_p \hat{x} - d_x \hat{p})}. \quad (2.11)$$

Phase-Space Representation

The states of a CV system are the set of positive density operators (ρ) on the Hilbert space \mathcal{H} . Instead of referring to states, one can refer to functions defined on the phase-space by analogy with classical dynamics. For later purpose it is most convenient to introduce the characteristic function which is the Fourier transform of the Wigner function. Using the Weyl operator D_ξ we define the (Wigner-)characteristic function as

$$\chi_\rho(\xi) = \text{Tr}[\rho D_\xi]. \quad (2.12)$$

The vector ξ belong to the real $2N$ -dimensional space, called phase-space. Each characteristic function is uniquely associated with a state, and they are related with each other via a Fourier-Weyl relation. The state ρ can be obtained from its characteristic function according to

$$\rho = \frac{1}{(2\pi)^N} \int d^{2N} \xi \chi_\rho(-\xi) D_\xi. \quad (2.13)$$

In turn, the quasi-probability distribution Wigner function as commonly used in quantum optics is related to the characteristic function via a Fourier transform,

$$W(\xi) = \frac{1}{(2\pi)^N} \int d^{2N} \zeta e^{i\xi^T \Omega \zeta} \chi_\rho(\zeta). \quad (2.14)$$

Wigner Function

The Wigner function can be written as follows in term of the eigenvectors of the quadrature operators,

$$W(x, p) = \frac{1}{(2\pi)^N} \int_{\mathbb{R}^N} \langle x - x' | \rho | x + x' \rangle e^{ix' \cdot p} d^N x', \quad x, p \in \mathbb{R}^N. \quad (2.15)$$

From an operational point of view, the Wigner function admits a clear interpretation in terms of homodyne measurement, as the marginal integral of the Wigner function over the variables $x_1, \dots, x_{N-1}, p_1, \dots, p_N$

$$p(x_N) = \int_{\mathbb{R}^{2N-1}} W(x, p) d^N p dx_1 \dots dx_{N-1}, \quad (2.16)$$

gives the probability of the result of homodyne detection on the remaining quadrature x_N .

Properties

Linearity The equation (2.15) being linear in ρ we deduce that the Wigner function W_ρ of a mixture of quantum states $\rho = \lambda\rho_1 + (1 - \lambda)\rho_2$ is the balanced average of the corresponding Wigner functions,

$$W_\rho(r) = \lambda W_{\rho_1}(r) + (1 - \lambda)W_{\rho_2}(r). \quad (2.17)$$

Wigner function of operators The Wigner function $W(r)$ is deeply related to the symmetrically ordered expressions of operators. More precisely, if the operator \hat{A} can be expressed as $\hat{A} = f(x, p)$ for $i = 1, \dots, N$, where f is a symmetrically ordered function of the quadratures operators (x, p) , then one can define a Wigner function $W_A(x, p) = f(x, p)/(2\pi)^N$ such that,

$$\text{Tr}[\rho\hat{A}] = (2\pi)^N \int_{\mathbb{R}^{2N}} W_\rho(r)W_A(r)d^{2N}r. \quad (2.18)$$

Let's consider some examples:

1. If the operator \hat{A} is just the identity $\hat{A} = \mathbb{I}_N$, his Wigner function reads $W_{\mathbb{I}_N} = 1/(2\pi)^N$ giving

$$\text{Tr}\rho = 1 = \int_{\mathbb{R}^{2N}} W_\rho(r)d^{2N}r. \quad (2.19)$$

2. If $\hat{A} = \rho$, we obtain the purity

$$\mu = \text{Tr}\rho^2 = (2\pi)^N \int_{\mathbb{R}^{2N}} W_\rho^2(x, p)d^N r. \quad (2.20)$$

3. If \hat{A} is a given pure state $|\Psi\rangle$ ($\hat{A} = |\Psi\rangle\langle\Psi|$), we obtain the fidelity F between the pure state $|\Psi\rangle\langle\Psi|$ and ρ ,

$$F = \text{Tr}[|\Psi\rangle\langle\Psi|\rho] = (2\pi)^N \int_{\mathbb{R}^{2N}} W_{|\Psi\rangle\langle\Psi|}(r)W_\rho(r)d^{2N}r. \quad (2.21)$$

2.2 Gaussian States

Gaussian states are defined through their properties that the characteristic function is a Gaussian function in phase-space. For a general density operator ρ we define the displacement vector ($d \in \mathbb{R}^{2N}$)

$$d = \langle\hat{r}\rangle = \text{Tr}[\rho\hat{r}], \quad (2.22)$$

and the positive-semidefinite symmetric $2N \times 2N$ covariance matrix γ :

$$\gamma_{ij} = \text{Tr}[\rho\{(\hat{r}_i - d_i), (\hat{r}_j - d_j)\}], \quad (2.23)$$

where $\{\}$ denotes the anticommutator. The Gaussian states are defined by a Gaussian characteristic function

$$\chi_\rho(\xi) = e^{-\frac{1}{4}\xi^T\Gamma\xi + iD^T\xi}, \quad (2.24)$$

which are characterized by $D = \Omega d$ and covariance matrix $\Gamma = \Omega\gamma\Omega$. Despite the infinite dimension of the Hilbert space in which it lives, a complete description of a Gaussian state of N modes requires only a polynomial number of real parameters for its full description. The Wigner function of Gaussian states reads,

$$W(r) = \frac{1}{\pi^{2N}\sqrt{\det\gamma}}e^{-(r-d)^T\gamma^{-1}(r-d)}. \quad (2.25)$$

Clearly, not all real symmetric $2N \times 2N$ matrix can be a legitimate covariance of a quantum state as the states must respect the Heisenberg uncertainty relation. If one requires the positive-semidefiniteness of the density operator ρ together with the CCR relation we obtain the condition

$$\gamma + i\Omega \geq 0, \quad (2.26)$$

which is the only necessary and sufficient constraint γ has to fulfill to be the covariance matrix of a physical Gaussian state, more generally it is also a necessary condition (but not sufficient) for non-Gaussian states. The constraint (2.26) generalizes the expression of Heisenberg uncertainty principle. Notice that all the states of the electromagnetic field presented in the previous chapter except the photon number states are indeed Gaussian states.

One mode Gaussian States

In the following we review the previously presented states:

Vacuum and Coherent states The vacuum state is a state centered at the origin of the phase space ($d = (0, 0)$) with a covariance matrix $\gamma = \mathbb{I}$. Coherent states being a displaced vacuum state we conclude that its covariance matrix is also the identity with a non null displacement vector $d = (d_x, d_p)$.

Squeezed State The squeezed vacuum state has null mean value as the vacuum state. His covariance matrix reads

$$\gamma = \begin{bmatrix} e^{-2r} & 0 \\ 0 & e^{2r} \end{bmatrix}, \quad (2.27)$$

where we observe that the uncertainty among one quadrature is squeezed (x if $r > 0$ and p if $r < 0$) and antisqueezed among the conjugate one. Squeezed coherent states have exactly the same covariance matrix but with a non null displacement. One can generate squeezed states along quadrature \hat{x}_θ by applying a previous phase shift θ before the squeezing operation.

Thermal State The thermal state has null mean value and covariance matrix

$$\gamma = \begin{bmatrix} V & 0 \\ 0 & V \end{bmatrix}. \quad (2.28)$$

The quantity V can be expressed in terms of the average number of photons \bar{n} contained in the thermal state as $V = 2\bar{n} + 1$. The vacuum can be seen as member of this class which contains no photons at all ($\bar{n} = 0$).

Noisy Coherent State One can generalize the thermal state by applying a displacement d to it. One then obtain a noisy version of the coherent state.

Generalization One mode Gaussian states would be completely characterized by the displacement operator $d = (d_x, d_p)$ and a 2×2 covariance matrix

$$\gamma = \begin{bmatrix} a & c \\ c & b \end{bmatrix}. \quad (2.29)$$

One can show that an arbitrary single mode Gaussian state can be generated from a thermal state by applying a squeezing operation and a subsequent rotation as we will show later in this chapter.

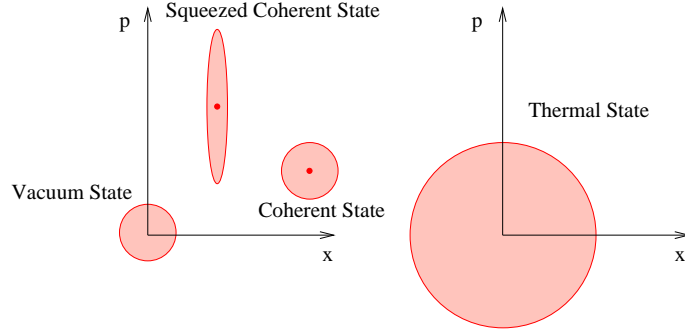


Figure 2.1: Phase space representation of most relevant one-mode Gaussian states.

Two modes Gaussian States

A general two mode Gaussian state is characterized by a mean $d = d_1 \otimes d_2$ and a covariance matrix

$$\gamma_{12} = \begin{bmatrix} \gamma_1 & C \\ C & \gamma_2 \end{bmatrix}. \quad (2.30)$$

where $\gamma_{1(2)}$ are the covariance matrix of the first and second mode after tracing the companion and C is the matrix that gives the correlation between the two modes, which can be either classical or quantum (entanglement).

Tensor Product State The case where $C = 0$ correspond to a tensor products of one mode Gaussian states

$$\gamma_{12} = \gamma_1 \oplus \gamma_2. \quad (2.31)$$

Two-Mode Squeezed State The two-mode squeezed vacuum state is a key resource for practical implementation of CV quantum information protocols such as teleportation, dense coding and quantum key distribution, playing an equivalent role as Bell pairs $((|00\rangle + |11\rangle)/\sqrt{2})$ in qubit quantum information. Its mean is null and its covariance matrix reads,

$$\gamma_{EPR} = \begin{bmatrix} \cosh 2r\mathbb{I} & \sinh 2r\sigma_z \\ \sinh 2r\sigma_z & \cosh 2r\mathbb{I} \end{bmatrix}, \quad (2.32)$$

where

$$\mathbb{I} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}. \quad (2.33)$$

Notice that tracing one mode leaves the other mode in a thermal state of variance $\cosh(2r) = 2\bar{n} + 1$.

As for the one mode case one can transform any two-mode Gaussian state in a two-mode thermal state by applying a given sequence of operations, as we will describe later in this chapter.

Multipartite Gaussian States

One can generalize the previous definitions to systems of N modes. The situation becomes highly non trivial if we separate the different modes among different partner which are only allowed to apply local operations and classical communication. This is the usual situation we encounter when we have entanglement distributed among different locations or in quantum key distribution analysis. The problem is highly complex and is at the moment a hot topic of research in quantum information.

Bipartite State The simpler case consist in a bipartite Gaussian state, where N modes are distributed to Alice and M modes to Bob. In this case the $N \times M$ modes covariance matrices reads,

$$\gamma_{AB} = \begin{bmatrix} \gamma_A & C \\ C & \gamma_B \end{bmatrix}. \quad (2.34)$$

where $\gamma_{A(B)}$ are the local covariance matrix and C is the correlations matrix between Alice and Bob states.

2.3 Gaussian Operations

The most straightforward definition of Gaussian operations states that an operation is Gaussian if it maps every Gaussian input state onto a Gaussian output state. In this section we will show that they exactly correspond to those operations that can be implemented by means of optical elements such as displacements, beamsplitters, phase-shifters, squeezers together with homodyne measurement. All Gaussian operations are experimentally accessible with present technology.

Gaussian Unitary Operations

The unitary operations preserving the Gaussian character of the states on which they act are generated by the set of the displacements (Weyl operators) D_ξ which correspond to Hamiltonians linear in the field operators and the quadratic Hamiltonian,

$$H = \sum_{j,k=1}^{2N} g_{jk} (\hat{r}_j \hat{r}_k + \hat{r}_k \hat{r}_j) / 2. \quad (2.35)$$

Displacement Operator The effect of a displacement D_z operator on any Gaussian state is to translate his mean $d_{out} = d_{in} + z$ leaving invariant the covariance matrix. More generally a displacement leaves invariant the shape of the Wigner function of any

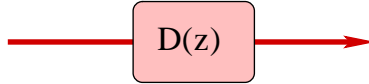


Figure 2.2: A displacement $D(z)$ operation acting on a single-mode field.

quantum state of light, included the non-Gaussian states, translating his mean-value.

Symplectic Transformations

As a consequence of the Stone-von Neumann theorem, any unitary transformation U_S generated from a quadratic Hamiltonian corresponds in phase-space to a symplectic operation (matrix) $S \in Sp(2N, \mathbb{R})$ which implements the mapping,

$$\hat{r}_{out} = S \hat{r}_{in} \quad (2.36)$$

and preserves the canonical commutation relation, which is the case if

$$S \Omega S^T = \Omega. \quad (2.37)$$

On the level of covariance matrix, a symplectic transformation S is reflected by a congruence

$$\gamma_{out} = S \gamma_{in} S^T. \quad (2.38)$$

Properties To any symplectic transformation S also S^T , S^{-1} , $-S$ are symplectic. Using $\Omega^T \Omega = \mathbb{I}$ (which implies $\Omega^T = \Omega^{-1} = -\Omega$) and $S\Omega S^T = \Omega$ we can prove $S^{-1} = -\Omega S^T \Omega$. The determinant of every symplectic matrix is $\det S = 1$.

Passive Transformations

A particularly important set of symplectic transformations is formed by those $S \in Sp(2N, \mathbb{R})$ that are moreover orthogonal, i.e., $P(N) = Sp(2N, \mathbb{R}) \cap O(2N)$. Those operations correspond with the passive transformations, such as phase shifts and beam-splitters, which preserve the number of photons.

Phase Shift A phase shift is a single-mode operation characterized by an angle (θ) which is equivalent to a rotation of the phase-space,

$$\begin{bmatrix} x \\ p \end{bmatrix}_{out} = S_{PS}(\theta) \bar{r}_{in} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ p \end{bmatrix}_{in}. \quad (2.39)$$

Beamsplitter The beamsplitters operation of transmittance T makes a coherent combination of two modes,

$$\begin{bmatrix} \bar{r}_1 \\ \bar{r}_2 \end{bmatrix}_{out} = S_{BS}(T) \bar{r} = \begin{bmatrix} \sqrt{T} \mathbb{I} & \sqrt{1-T} \mathbb{I} \\ -\sqrt{1-T} \mathbb{I} & \sqrt{T} \mathbb{I} \end{bmatrix} \begin{bmatrix} \bar{r}_1 \\ \bar{r}_2 \end{bmatrix}_{in}. \quad (2.40)$$

Any passive transformation over N modes can be decomposed in a network of beamsplitters and phase shifts. Such transformation does not change the eigenvalues of the covariance matrix which is equivalent to preserving the total number of photons.

Active Transformations

The set of symplectic transformations which are not passive $A(N) = S \in \{Sp(2N, \mathbb{R}) \setminus P(N)\}$ are called active transformations as they inject photons on the system, usually achieved by pumping a nonlinear media.

Squeezing The most important active transformation is the squeezed operation presented in the previous chapter, obtained by Optical Parametric Amplification (OPA) on a nonlinear media pumped by high intensity source. The symplectic transformation of a squeezing operation reads,

$$\begin{bmatrix} x \\ p \end{bmatrix}_{out} = S_{Sq}(s) \bar{r} = \begin{bmatrix} e^{-s} & 0 \\ 0 & e^s \end{bmatrix} \begin{bmatrix} x \\ p \end{bmatrix}_{in}, \quad (2.41)$$

where s is the squeezing factor, $s > 0$ giving squeezing among x and $s < 0$ among p . In order to squeeze among any other quadrature one has to combine a phase rotation with a squeezing operation $S_{Sq}(s)S_{PS}(\theta)$.

General Symplectic Transformation Combining passive operations with squeezing is enough to generate any symplectic transformation (S) over N modes, as shown in Fig. 2.3. Using the Bloch-Messiah reduction theorem one can decompose any S symplectic transformation into a linear interferometer K , a parallel set of single-mode squeezers $S(s_i)$ and a second linear interferometer L ,

$$S = K \bigoplus_{j=1}^N \begin{bmatrix} e^{-s_j} & 0 \\ 0 & e^{s_j} \end{bmatrix} L, \quad (2.42)$$

with $K, L \in P(N)$ and $s_j \in \mathbb{R}$, as shown in [33].

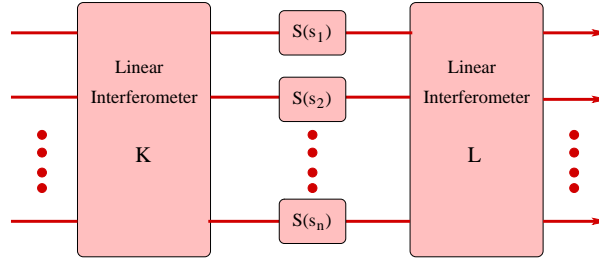


Figure 2.3: An arbitrary symplectic transformation can be decomposed into a linear interferometer K , a parallel set of single-mode squeezers $S(s_i)$ and a second linear interferometer L .

Two Important Examples

Even if all the symplectic transformations can be decomposed using a Bloch-Messiah decomposition, we are going to present two different active two-mode operations which are relevant from the experimental point of view, the two-mode squeezer and the QND-measurement (Quantum Non Demolition) interaction.

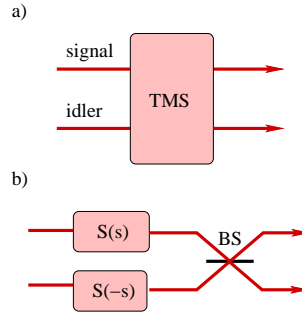


Figure 2.4: Two different ways of implementing a two-mode squeezing operation over two modes, signal and idler: a) spatially and temporally matching the signal, idler and pump on a NOPA . b) using two OPA to generate two squeezed states among orthogonal directions and combining them on a balanced beamsplitter.

Two-Mode Squeezer Pumping a NOPA realize a two-mode squeezing operation over the signal and idler modes,

$$\begin{bmatrix} \bar{r}_1 \\ \bar{r}_2 \end{bmatrix}_{out} = S_{BS}(r)\bar{r} = \begin{bmatrix} \cosh r\mathbb{I} & \sinh r\sigma_z \\ \sinh r\sigma_z & \cosh r\mathbb{I} \end{bmatrix} \begin{bmatrix} \bar{r}_1 \\ \bar{r}_2 \end{bmatrix}_{in}. \quad (2.43)$$

This operation is extremely important as it is the usual way of generating entanglement in continuous variable quantum information experiments. The Bloch-Messiah decomposition corresponds to two squeezing operations over orthogonal directions which are subsequently combined into a balanced beamsplitter, as shown in Fig. 2.4.

QND-measurement The CV QND gate realizes a controlled displacement over the x quadrature of the target mode depending on the value of the x quadrature of the control mode. In order to be unitary the operation realizes a similar operation on the

conjugate quadrature where the roles (target and control) are inverted and the sign of the displacement changed,

$$\begin{bmatrix} x_1 \\ p_1 \\ x_2 \\ p_2 \end{bmatrix}_{out} = S_{BS}(\kappa)\bar{r} = \begin{bmatrix} 1 & 0 & \kappa & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & -\kappa & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ p_1 \\ x_2 \\ p_2 \end{bmatrix}_{in}. \quad (2.44)$$

This operation is also extremely important from the experimental point of view as it is the interaction that takes place between polarized light and the total spin of atomic ensembles [62] which is currently used in quantum memory experiments [114]. In the particular case $\kappa = 1$ we call it also the continuous-variable C-NOT gate, as it generalizes the C-NOT gate for qubits.

Local Operation

In a multipartite scenario where the N modes are distributed among different locations an important class of operations are the local symplectic operations $Sp(2, \mathbb{R})^{\otimes N}$ as they correspond, on the Hilbert space level, to tensor product of Gaussian unitary operations $U^{\otimes N}$. Those correspond to the set of operations that are locally accessible without using quantum communication between the partners. It is trivial to show that the determinant of the submatrices $\gamma_{A(B)}$ and C of a bipartite state, i.e. eq. (2.34), are invariant under local symplectic operations, as $S = S_1 \otimes S_2$ and $\det S_1 = \det S_2 = 1$, showing that the correlations among the parties cannot be altered by local operations, as expected.

Standard Form Any two-mode covariance matrix

$$\gamma_{AB} = \begin{bmatrix} \gamma_A & C \\ C^T & \gamma_B \end{bmatrix} \quad (2.45)$$

can be transformed by local Gaussian operations into a standard form

$$\gamma_{AB} = \begin{bmatrix} a & 0 & c_+ & 0 \\ 0 & a & 0 & c_- \\ c_+ & 0 & b & 0 \\ 0 & c_- & 0 & b \end{bmatrix}. \quad (2.46)$$

The idea is first to apply a round of local rotations $S_A(\theta) \oplus S_B(\xi)$ in order to diagonalize the diagonal submatrices $\gamma_{A(B)}$. Then we apply a local squeezing operations $S_A(s_a) \oplus S_B(s_b)$ in order to transform the diagonal blocs to $\gamma_A = a\mathbb{I}$ and $\gamma_B = b\mathbb{I}$. Finally we apply a second round of local rotations $S_A(\theta') \oplus S_B(\xi')$ that diagonalize the correlation submatrix C' ($C' = S_A(s_a)S_A(\theta)CS_B^T(\xi)S_B^T(s_b)$), as shown in [63].

CP Maps

The set of unitary operations does not contain all the operations that can be applied to a general quantum system. The unitary operation being reversible one has to add the class of irreversible operations, the so-called Completely Positive maps (CP maps), in order to have a complete description, see Appendix A. The Gaussian CP maps are completely defined by two matrices X and Y such that the covariance matrix of the final N -mode state reads,

$$\gamma_{out} = X\gamma_{in}X^T + Y \quad (2.47)$$

where X and Y are $2N \times 2N$ matrices and Y is symmetric [72]. The quantum positivity of the quantum operation is reflected by the condition

$$Y + i\Omega - iX\Omega X^T \geq 0. \quad (2.48)$$

The mean value (d) transformation reads,

$$d_{out} = X d_{in}. \quad (2.49)$$

Partial Trace

Consider a bipartite quantum state ρ_{AB} . In the phase-space representation tracing mode B is equivalent to integrating the Wigner function among the quadratures of the traced mode (x_B, p_B) , as shown in (2.18). One can show that this is equivalent to setting $\xi_B = 0$ at the characteristic function

$$\chi_A(\xi_A) = \text{Tr}_A[\rho_A D_{\xi_A}] = \text{Tr}_{AB}[\rho_{AB} D_{\xi_A} \otimes \mathbb{I}_B] = \chi_A(\xi_A, \xi_B = 0). \quad (2.50)$$

It is then trivial to see that for Gaussian states the partial trace (PT) of mode B gives as output a Gaussian state of covariance matrix γ_A , where γ_A is the diagonal bloc corresponding to mode A of the initial bipartite covariance matrix γ_{AB} .

$$\gamma_{AB} = \begin{bmatrix} \gamma_A & C \\ C^T & \gamma_B \end{bmatrix} \xrightarrow{PT} \gamma_A. \quad (2.51)$$

Similarly the first moment of the output state is the part of the mean corresponding to mode A ,

$$d_{AB} = (d_A, d_B) \xrightarrow{PT} d_A. \quad (2.52)$$

Gaussian Channels

Lossy Channel The lossy channel of transmittance T is characterized by $X = \sqrt{T}\mathbb{I}$ and $Y = (1 - T)\mathbb{I}$. It can be modeled by a beamsplitter of transmittance T , as shown in Fig. 2.5

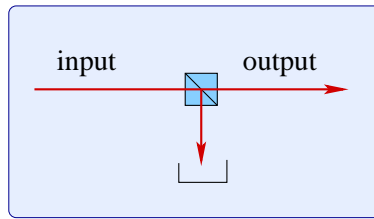


Figure 2.5: A lossy channel can be modeled by placing a beamsplitter of transmittance T .

Thermal Noise Channel The thermal noise channel of transmittance T and excess noise ϵ is characterized by $X = \sqrt{T}\mathbb{I}$ and $Y = T\chi\mathbb{I}$, where χ is the added noise referred to the input

$$\chi = \frac{1 - T}{T} + \epsilon. \quad (2.53)$$

It can be modeled by combining a thermal state of variance $N = T\chi/(1 - T)$ with the input signal at a beamsplitter of transmittance T , as shown in Fig. 2.6.

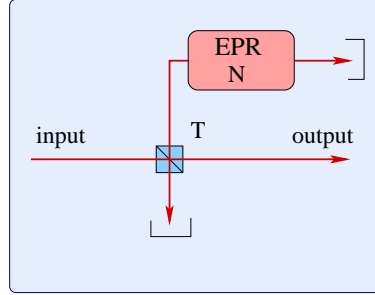


Figure 2.6: A thermal noise channel can be modeled by combining a thermal state of variance $N = T\chi/(1 - T)$ with the input signal at a beamsplitter of transmittance T .

Amplification Channel The amplification channel of amplification factor $\eta \geq 0$ is characterized by $X = \sqrt{\eta}\mathbb{I}$ and $Y = (\eta - 1)\mathbb{I}$. It can be modeled by injecting the input signal into a two-mode squeezer (Fig. 2.4), with squeezing factor such that $\eta = \cosh^2 r$, and finally tracing the idler mode.

Classical Noise Channel The classical noise channel adds noise of variance V to a quantum state, with $X = \mathbb{I}$ and $Y = V\mathbb{I}$. It can be modeled by applying two C-NOT operations, with an interaction parameter $\kappa = V$, between the signal mode and two ancillary vacuums as shown in Fig. 2.7.

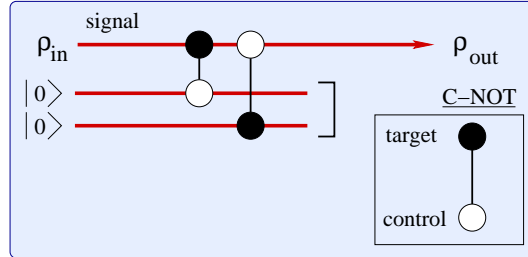


Figure 2.7: A classical noise channel can be modeled by applying two C-NOT operations, with an interaction parameter $\kappa = V$, over the signal mode. The first C-NOT generates the noise over the x quadrature and the second over p .

Measurement

Homodyne Measurement

The outcome of a projective measurement M is given by $\text{Tr}[\rho M]$ which can be calculated using equation (2.18). As explained in subsection 2.1 the probability distribution of an homodyne measurement of the x quadrature over a quantum mode is given by the marginal integral of the Wigner function over the p quadrature

$$p(x) = \int_{\mathbb{R}} W(x, p) dp. \quad (2.54)$$

Measuring over a general quadrature $\hat{x}_\theta = \cos \theta \hat{p} + \sin \theta \hat{q}$ the probability distribution of the measurement can be calculated applying a virtual rotation of $-\theta$ and integrating

over p ,

$$p(x) = \int_{\mathbb{R}} W(S^{-1}(\theta)(x, p)) dp. \quad (2.55)$$

Both equations can be trivially extended to measurement over multiple modes.

Partial Measurement

In the following we are going to consider partial measurements over a bipartite system. Given a Gaussian bipartite state ρ_{AB} with covariance matrix

$$\gamma_{AB} = \begin{bmatrix} \gamma_A & C \\ C^T & \gamma_B \end{bmatrix} \quad (2.56)$$

and mean $d_{AB}^{in} = (d_A^{in}, d_B^{in})$. In appendix B we calculated the final state of system A after projecting mode B into a given quantum state of covariance matrix γ_M and mean m .

Homodyne measurement The case of homodyne measurement is special as it corresponds to the case $\gamma_M = \lim_{r \rightarrow \infty} (e^{2r}, e^{-2r})$, then the inverse in equations (B.5) and (B.5) is to be understood as the inverse on the range. The homodyne measurement gives [72],

$$\gamma_A^{out} = \gamma_A - C(X\gamma_B X)^{MP} C^T, \quad (2.57)$$

where $X = \text{diag}(1, 0, 1, 0, \dots, 1, 0)$ and MP denotes the inverse on the range. The mean value reads,

$$d_A^{out} = C(X\gamma_B X)^{MP} (m - d_B^{in}) + d_A^{in}, \quad (2.58)$$

where $m = (X_1, 0, X_2, 0, \dots, X_N, 0)$, X_i being the result of the homodyne measurement on mode B_i . A measurement over any other quadrature can be obtained by applying a phase rotation before the homodyne measurement. Remark that the covariance matrix γ_A^{out} does not depend on the result of the measurement m , which is a very important property of Gaussian state.

Heterodyne measurement Heterodyne measurement can be decomposed in two different homodyne measurements preceded by a balanced beamsplitter, so its effect can be determined using the previous result for homodyning and adding ancillary systems. Equivalently it can be calculated using directly the result of appendix B and noticing that an heterodyne detection is a projection over coherent states ($\gamma_M = \mathbb{I}$). After an heterodyne measurement on mode B the covariance matrix of mode A reads,

$$\gamma_A^{out} = \gamma_A - C(\gamma_B + \mathbb{I})^{-1} C^T, \quad (2.59)$$

and the mean value reads,

$$d_A^{out} = \sqrt{2}C(\gamma_B + \mathbb{I})^{-1} (m - d_B^{in}) + d_A^{in}, \quad (2.60)$$

where $m = (X_1, P_1, X_2, P_2, \dots, X_N, P_N)$ is the result of the heterodyne measurement on B .

Projection on Vacuum The vacuum being just a coherent state of null mean value, projecting over the vacuum on mode B corresponds to the heterodyne measurement of output $m = 0$. The difference with heterodyning is that the projection on vacuum is a probabilistic operation. The projection can be made deterministic by applying an heterodyne measurement over B and then applying a displacement ($d = -\sqrt{2}C(\gamma_B + \mathbb{I})^{-1}(m - d_B^{in})$) over A in order to restore the null mean.

2.4 Normal Modes Decomposition

Thanks to Williamson theorem [183] we know that for any N -mode covariance matrix γ there is a symplectic transformation S that performs a symplectic diagonalization

$$S\gamma S^T = \nu, \quad (2.61)$$

where ν is a tensor product of thermal states, called the Williamson normal form,

$$\nu = \bigoplus_{k=1}^N \begin{bmatrix} \nu_k & 0 \\ 0 & \nu_k \end{bmatrix}. \quad (2.62)$$

The symplectic eigenvalues ν_k being the eigenvalues of the matrix $|i\Omega\gamma|$, where $|A|$ stands for $\sqrt{A^\dagger A}$.

States with null mean value For a null mean value state the Williamson decomposition gives a way of transform the Gaussian state into a product of thermal states. This symplectic diagonalization over the phase-space plays a similar role as the diagonalization of the density operator on the Hilbert space.

States with no null mean value Gaussian states of covariance matrix γ and mean d have also a Williamson decomposition, which is based on tensor product of displaced thermal states of covariance matrix ν_k and mean l_k . We first decompose the covariance matrix γ into a product of tensor thermal states $\bigotimes_k \nu_k$ by selecting the proper symplectic transformation S , as before. Finally, once S is known, the displacement vector $l = (l_1, \dots, l_N)$ of the thermal states is calculated by solving the following linear system,

$$l = S^{-1}d. \quad (2.63)$$

Uncertainty Relation The uncertainty relation $\gamma + i\Omega \geq 0$ is equivalent to $\nu + i\Omega \geq 0$ which in term of symplectic eigenvalues read

$$\nu_i \geq 1 \quad \forall i = 1, \dots, n. \quad (2.64)$$

Purity The symplectic transformations being an unitary operation it is trivial to see that a state is pure if and only if $\nu = \mathbb{I}$. More precisely, the purity μ of a Gaussian state ρ of covariance matrix γ reads,

$$\mu = \text{Tr}\rho^2 = \frac{1}{\sqrt{\det \gamma}}. \quad (2.65)$$

The determinant is then a symplectic invariant, as $\det S = 1$, which gives,

$$\det \gamma = \det \nu = \prod_{i=1}^N \nu_i^2. \quad (2.66)$$

One-Mode Normal Decomposition

The determinant of the covariance matrix γ_1 being an invariant over symplectic transformations, as pointed previously, it is easy to derive the normal decomposition of one mode as $\nu_1^2 = \det \gamma_1$.

Two-Mode Normal Decomposition

In order to find the symplectic eigenvalues $\nu_{1,2}$ of the two-mode covariance matrix

$$\gamma_{12} = \begin{bmatrix} \gamma_1 & C_{1-2} \\ C_{1-2}^T & \gamma_2 \end{bmatrix}, \quad (2.67)$$

we need to define a second symplectic invariant,

$$\Delta = \nu_1^2 + \nu_2^2 = \det \gamma_1 + \det \gamma_2 + 2 \det C_{1-2}, \quad (2.68)$$

where the first reads,

$$\det \gamma_{12} = \nu_1^2 \nu_2^2. \quad (2.69)$$

It is easy to see that the symplectic eigenvalues are solutions of the second order polynomial

$$z^2 - \Delta z + \det \gamma_{12} = 0, \quad (2.70)$$

which gives the solution,

$$\nu_{1,2}^2 = \frac{1}{2} \left[\Delta \pm \sqrt{\Delta^2 - 4 \det \gamma_{12}} \right]. \quad (2.71)$$

Three Modes and Generalization

In order to find the symplectic eigenvalues $\nu_{1,2,3}$ of the three-mode covariance matrix

$$\gamma_{123} = \begin{bmatrix} \gamma_1 & C_{1-2} & C_{1-3} \\ C_{1-2}^T & \gamma_2 & C_{2-3} \\ C_{1-3}^T & C_{2-3}^T & \gamma_3 \end{bmatrix}, \quad (2.72)$$

we need to define a three symplectic invariants,

$$\Delta_1^3 = \nu_1^2 + \nu_2^2 + \nu_3^2, \quad (2.73)$$

$$\Delta_2^3 = \nu_1^2 \nu_2^2 + \nu_2^2 \nu_3^2 + \nu_1^2 \nu_3^2, \quad (2.74)$$

$$\Delta_3^3 = \nu_1^2 \nu_2^2 \nu_3^2, \quad (2.75)$$

which can be calculated from the covariance matrix with,

$$\begin{aligned} \Delta_1^3 &= \det \gamma_1 + \det \gamma_2 + \det \gamma_3 + 2 \det C_{1-2} + 2 \det C_{1-3} \\ &\quad + 2 \det C_{2-3}, \\ \Delta_2^3 &= \det \gamma_{12} + \det \gamma_{23} + \det \gamma_{13} + 2 \det C_{12-23} + 2 \det C_{12-13} \\ &\quad + 2 \det C_{23-13}, \\ \Delta_3^3 &= \det \gamma_{123}, \end{aligned} \quad (2.76)$$

where the matrix C_{ij-kl} reads,

$$C_{ij-kl} = \begin{bmatrix} \alpha_{ik} & \alpha_{il} \\ \alpha_{jk} & \alpha_{jl} \end{bmatrix}, \quad (2.77)$$

where α_{mn} is an elementary submatrice describing the correlations between a pair of modes of the covariance matrix γ_{123} . The symplectic eigenvalues being the solutions of the third order polynomial

$$z^3 - \Delta_1^3 z^2 + \Delta_2^3 z - \Delta_3^3 = 0. \quad (2.78)$$

This technique can be generalized to N modes by defining new symplectic invariants as shown in [173, 174],

$$\Delta_j^N(\nu_1, \dots, \nu_N) = \sum_{\mathcal{S}_j^k} \prod_{k \in \mathcal{S}_j^k} \nu_k^2, \quad (2.79)$$

where the sum runs over all the possible j -subsets \mathcal{S}_j^k of the first N natural integers (over all the possible combinations of j integers). The symplectic eigenvalues can be calculated from the $2k \times 2k$ submatrices of γ obtained by selecting all the combinations of 2×2 blocks describing either one mode or the correlations between a pair of modes, generalizing 2.76, as show in [173].

2.5 Phase-Space Schmidt Decomposition

In general any quantum state ρ_A can be diagonalize by choosing a proper unitary operation U_A , such that

$$U_A \rho_A U_A^\dagger = \sum_i \lambda_i |\psi_i\rangle \langle \psi_i|, \quad (2.80)$$

where $|\psi_j\rangle \langle \psi_j|$ are orthogonal eigenvectors of real eigenvalues λ_i satisfying $\sum_i \lambda_i^2 = 1$. One can then introduce another system, denoted R (reference system), and define a bipartite pure state Ψ_{RA} such that $\rho_A = \text{Tr}_R[|\Psi\rangle \langle \Psi|_{AR}]$, see Appendix A. The procedure is to define a pure state whose Schmidt basis [166] is just the basis for which the mixed state is diagonal, with the Schmidt coefficients being $\sqrt{\lambda_i}$,

$$|\Psi\rangle_{RA} = \sum_i \sqrt{\lambda_i} |i\rangle_R |\psi_i\rangle_A. \quad (2.81)$$

Similarly we know that for a Gaussian state with null mean value and covariance matrix γ_A there is a symplectic transformation S that transforms the covariance matrix γ_A into a tensor product of thermal states $\otimes_k \nu_k \mathbb{I}$ as shown in section 2.4. Then using the properties of symplectic operations of subsection 2.3 we can write,

$$\gamma_A = (\Omega S^T) \left[\bigoplus_{k=1}^N \nu_k \mathbb{I} \right] (\Omega S^T)^T. \quad (2.82)$$

Each thermal state (ν_k) could be seen as resulting from tracing half of an EPR pair ($\gamma_{EPR}^{\nu_k}$) (with squeezing factor $\cosh 2r_k = \nu_k$). Thus, applying the symplectic operation $(S\Omega)_A \otimes (S\Omega)_R$ over the ensemble of EPR states $\otimes_k \gamma_{EPR}^{\nu_k}$ we obtain a valid purification of γ_A which reads,

$$\gamma_{AR} = \begin{bmatrix} \gamma_A & C \\ C^T & \gamma_R \end{bmatrix}, \quad (2.83)$$

with

$$\gamma_R = \gamma_A \quad \text{and} \quad C = (\Omega S^T) \left[\bigoplus_{k=1}^N \sqrt{\nu_k^2 - 1} \sigma_z \right] (\Omega S^T)^T, \quad (2.84)$$

where $\sigma_z = \text{diag}(1, -1)$. The mean of the reference system being just equal to that of A , $d_R = d_A$.

2.6 Teleportation and Cloning

In the following we are going to present two relevant examples of quantum information processing with continuous variables that we will use later in this thesis, teleportation and cloning of continuous variables.

Generalization of Bell Basis and Measurement

In qubit based quantum information the maximally entangled states are the four Bell states $|\Phi^\pm\rangle = [|00\rangle \pm |11\rangle]/\sqrt{2}$ and $|\Psi^\pm\rangle = [|01\rangle \pm |10\rangle]/\sqrt{2}$. One can generate any of the four Bell states by injecting $[|0\rangle \pm |1\rangle]/\sqrt{2}$ on the control mode of a C-NOT gate and $\{|0\rangle, |1\rangle\}$ on the target state, as shown in [136]. It is trivial to see that the Bell measurement consists in a C-NOT operation followed by a measurement on the basis $[|0\rangle \pm |1\rangle]/\sqrt{2}$ ($\{|0\rangle, |1\rangle\}$) on the outgoing control (target) modes.

One can generalize this Bell basis to continuous variable by using the following analogy. The basis $\{|0\rangle, |1\rangle\}$ translates into x -squeezed state displaced among the x quadrature, where the basis $\{[|0\rangle \pm |1\rangle]/\sqrt{2}\}$ translates into p -squeezed state displaced along the p quadrature. Then the C-NOT gate is replaced by a beamsplitter as we know that injecting two conjugate squeezed vacuum states into a beamsplitter outputs an EPR state. The difference is that now the mean of the EPR is non null ($d = (d_x, d_p, d_x, -d_p)\sqrt{2}$). It is trivial to see that the equivalent Bell measurement in order to estimate d consists in mixing the two modes on a balanced beamsplitter and measuring x on one output and p on the other using homodyne measurement.

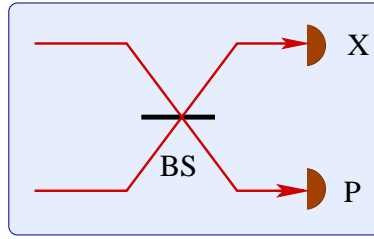


Figure 2.8: Generalized Bell measurement to continuous variables.

Teleportation

The proposal of teleportation of continuous variables [35] and its experimental demonstration [83] in 1998 started the field of quantum information with continuous variables. In order to unconditionally teleport a given quantum state of the electromagnetic field from Alice to Bob, we need first to distribute an entangled EPR pair between Alice and Bob, as shown in Fig. 2.9. Once Alice and Bob share an EPR pair (two-mode squeezed vacuum of variance $\cosh 2r$) Alice applies a Bell measurement by combining the input mode ρ_{in} and her half of the EPR pair into a balanced beamsplitter and measures the x -quadrature on one output and p on the other. Then she communicates the measurement result, using a classical channel, to Bob who applies a displacement operation on half of his EPR pair proportional (gain g) to Alice's measurement result (\hat{x}_m, \hat{x}_m) , as shown in Fig. 2.9. The advantage of the teleportation using continuous variables is that it is unconditional as the Bell measurement can be implemented using linear optics and homodyne detection as opposed to qubit teleportation with photons where it is impossible to implement an unconditional C-NOT gate using linear optics. The disadvantage is that the efficiency of the operation is upperbounded by the amount of squeezing used. In order to achieve perfect efficiency ($\rho_{out} = \rho_{in}$) we need an infinite energy EPR state which is impossible to generate in practice. In a realistic scenario even with perfect quantum channels, the output will always be a slightly noisy version of ρ_{in} , with the fidelity of the teleportation increasing with the squeezing parameter of the EPR pair, as we show bellow.

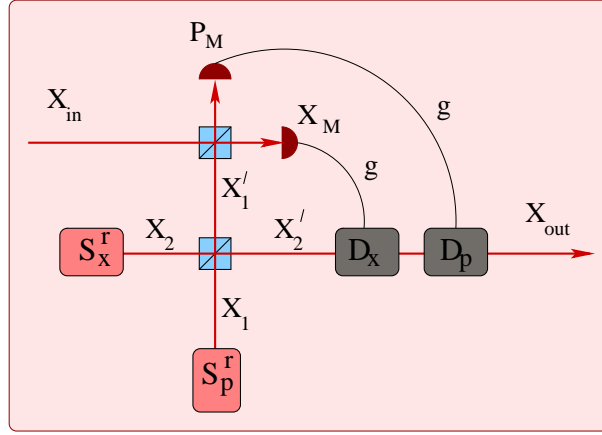


Figure 2.9: Teleportation scheme used in the experimental demonstration of unconditional teleportation [83]. Alice first generates an EPR pair by mixing a one x -squeezed vacuum state (X_2) with a p -squeezed vacuum state (X_1) in a balanced beamsplitter and send half of the EPR pair (X'_2) to Bob. Then, Alice applies a Bell measurement over the input mode X_{in} and X'_1 , that consist in an heterodyne measurement applied to X_{in} where the second input mode of the measurement is not the vacuum but mode X'_1 . Subsequently Alice communicates the measurement result (X_M, P_M) to Bob via a classical communication channel. Finally Bob depending on the measurement result and the gain g displaces mode X'_2 over x (D_x) and p (D_p).

Detailed Description of Teleportation

Because all the canonical transformations are symmetric on x and p quadratures we will just detail the operation for the quadrature x . Alice starts preparing two squeezed vacuum states, \hat{x}_2 squeezed over x and \hat{x}_1 squeezed over p .

$$\hat{x}_1 = e^r \hat{x}_1^{(0)}, \quad (2.85)$$

$$\hat{x}_2 = e^{-r} \hat{x}_2^{(0)}. \quad (2.86)$$

She subsequently mixes them on a balanced beamsplitter generating an EPR state,

$$\hat{x}'_1 = [e^{-r} \hat{x}_2^{(0)} - e^r \hat{x}_1^{(0)}] / \sqrt{2}, \quad (2.87)$$

$$\hat{x}'_2 = [e^{-r} \hat{x}_2^{(0)} + e^r \hat{x}_1^{(0)}] / \sqrt{2}. \quad (2.88)$$

After sending half of the EPR to Bob, Alice applies a Bell measurement by mixing \hat{x}'_1 and \hat{x}_{in} on a balanced beamsplitter and measuring x on one output and p on the other (Bell measurement),

$$\hat{x}_M = \frac{1}{\sqrt{2}}[\hat{x}_{in} + \hat{x}'_1] = \frac{1}{\sqrt{2}}\hat{x}_{in} + \frac{1}{2}[e^{-r} \hat{x}_2^{(0)} - e^r \hat{x}_1^{(0)}]. \quad (2.89)$$

Bob depending on the measurement result and the gain factor g displaces mode \hat{x}'_2 ,

$$\hat{x}_{out} = \hat{x}'_2 + g\hat{x}_M = \frac{g}{\sqrt{2}}\hat{x}_{in} + \frac{e^{-r}}{\sqrt{2}}\left[1 + \frac{g}{\sqrt{2}}\right]\hat{x}_2^{(0)} + \frac{e^r}{\sqrt{2}}\left[1 - \frac{g}{\sqrt{2}}\right]\hat{x}_1^{(0)}. \quad (2.90)$$

We see that by choosing $g = \sqrt{2}$ we obtain

$$\hat{x}_{out} = \hat{x}_{in} + \sqrt{2}e^{-r}\hat{x}_2^{(0)}, \quad (2.91)$$

which is a noisy version of the input quadrature as we have an added noise $\sqrt{2}e^{-r}\hat{x}_2^{(0)}$, which becomes negligible for high squeezing parameters.

Fidelity

It is easy to show, using (2.21) and (2.14), that for Gaussian states with null mean values the teleportation fidelity reads,

$$\text{Tr}[\rho_{out}|\Psi\rangle\langle\Psi|_{in}] = \int \chi(\xi)_{in}\chi(-\xi)_{out} = \frac{1}{\sqrt{\gamma_{in} + \gamma_{out}}}. \quad (2.92)$$

Coherent States The teleportation preserving the mean of the state, the fidelity of teleporting a coherent state will be the same as that of teleporting the vacuum which reads,

$$F = \frac{1}{1 + e^{-2r}}. \quad (2.93)$$

It is easy to see that there are two limiting implementations. Firstly, when the squeezing factor of the EPR (r) is null, the teleportation becomes a classical measurement and state preparation scheme with fidelity $F = 1/2$. This bound is very important as observing teleportation fidelities $\geq 1/2$ is necessary to claim a successful quantum teleportation experiment. Secondly, when the squeezing factor (r) is infinity, which needs an infinite source of energy, the teleportation succeeds perfectly ($F = 1$).

Entanglement Swapping

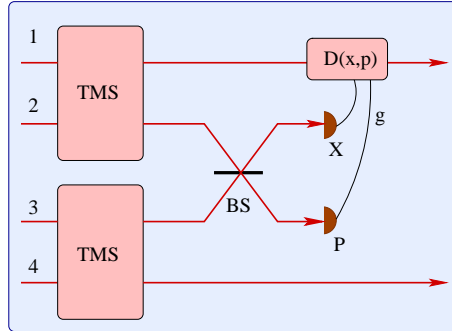


Figure 2.10: One can then entangle modes 1 and 4, which have never interacted before by applying entanglement swapping. In an intermediate station we apply a Bell measurement over two modes 2 and 3, coming from two different EPR pairs, and displace ($D(x, p)$) mode 1 (or 4) depending on the measurement results (x, p).

An ideal teleportation being equivalent to a quantum noiseless channel it preserves the coherence of the input state, for example it preserves the quantum superposition of two states. More interestingly if the input state is entangled with another system the teleportation will preserve the entanglement. One can then entangle two particles that have never interacted before by applying entanglement swapping, where we apply a teleportation to one mode of an EPR state, as show in Fig. 2.10.

Cloning Machine

Since the seminal work of Wootters and Zurek [199] it is well known that a machine, called *perfect cloning machine* (PCM) that outputs two perfect copies of a given quantum input state,

$$|\psi\rangle_{in} \xrightarrow{PCM} |\psi\rangle \otimes |\psi\rangle, \quad (2.94)$$

is forbidden by the laws of quantum mechanics. More precisely, perfect cloning is possible if and only if the input state ($|\psi\rangle_{in}$) is drawn from a set of orthogonal states, a simple von Neumann measurement will allow to identify it without disturbance and prepare as many copies as we want. On the contrary, when the input state is drawn from a set of non-orthogonal states perfect cloning becomes impossible.

Even if perfect cloning is forbidden, one can define approximate cloning machines, called *Quantum Cloning Machines* that output imperfect copies ($\rho_{1(2)}^\psi$) of the input state,

$$|\psi\rangle_{in} \xrightarrow{CM} \rho_1^\psi \otimes \rho_2^\psi. \quad (2.95)$$

For a given set of states one can then optimize the fidelity ($F = \langle \psi | \rho_i^\psi | \psi \rangle$) of the output copies among all the possible quantum copying machines, see [40, 165] for a review on quantum cloning.

No-Cloning Theorem and Quantum Cryptography The proof of the no-cloning theorem is at the core of quantum cryptography, one of the most important applications of quantum information. Wootters and Zurek result forbids a potential eavesdropper from copying a quantum state without disturbing it, allowing two trustfull partners to detect the action of an eavesdropper by monitoring the noise of their quantum communication. The study of the cloning machines is intimately related to quantum cryptography as the usual optimal attack on a given protocol is usually an asymmetric cloning machine which is a generalization of the previously presented machine where we allow different fidelities of the copies.

No Cloning Theorem for Continuous Variable

One can generalize the concept of cloning machine to the continuous variable framework, where an input mode ($\hat{x}_{in}, \hat{p}_{in}$) is cloned into two output noisy versions ($(\hat{x}_{1(2)}, \hat{x}_{1(2)})$), as shown in Fig. 2.11,

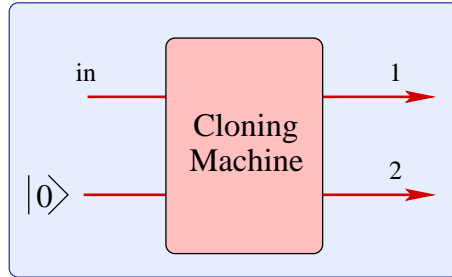


Figure 2.11: The cloning machine generates two noisy copies ($(\hat{x}_{1(2)}, \hat{x}_{1(2)})$) of the input state ($\hat{x}_{in}, \hat{x}_{in}$).

$$\hat{x}_{1(2)} = \hat{x}_{in} + \hat{x}_{1(2)}^N, \quad (2.96)$$

$$\hat{p}_{1(2)} = \hat{p}_{in} + \hat{p}_{1(2)}^N, \quad (2.97)$$

where $\hat{x}_{1(2)}^N$ stands for the added noise on the output modes. One can then derive a generalized uncertainty relation for the added noise (measured on shot-noise units) on the output modes [43],

$$\Delta \hat{x}_1^N \Delta \hat{p}_2^N \geq 1, \quad (2.98)$$

$$\Delta \hat{p}_1^N \Delta \hat{x}_2^N \geq 1, \quad (2.99)$$

which clearly shows that it is impossible to have two perfect copies and allows us at the same time to lowerbound the minimal disturbance of the cloning operation.

Linear Amplifier

A key element of the cloning of continuous variables is the use of the linear amplifier which is nothing else than the previously presented two-mode squeezer operation (2.43), where we feed the signal input (s) with the mode we want to amplify (\bar{r}_{in}) and the idler mode (i) with vacuum ($\bar{r}^{(0)}$),

$$\begin{bmatrix} \bar{r}_s \\ \bar{r}_{id} \end{bmatrix}_{out} = \begin{bmatrix} \sqrt{G} & \sqrt{G-1}\sigma_z \\ \sqrt{G-1}\sigma_z & \sqrt{G}\mathbb{I} \end{bmatrix} \begin{bmatrix} \bar{r}_{in} \\ \bar{r}_{id}^{(0)} \end{bmatrix}_{in}, \quad (2.100)$$

where the gain is proportional to the squeezing parameter ($G = \cosh 2r$).

Cloning of Coherent states

Using the cloning uncertainty relation it is trivial to see that for symmetric noise on both quadratures, the optimal copy of a coherent state has a unit of added shot noise. In Fig. 2.12 we present the implementation of the cloning machine presented in [34, 78]. First we fix the gain of the amplifier at $G = 2$ obtaining at the outputs

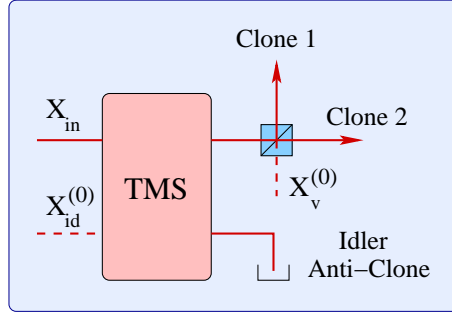


Figure 2.12: The cloning machine can be implemented by injecting the input mode \hat{x}_{in} in the signal input of a two-mode squeezer and dividing the signal amplified output in two clones using a balanced beamsplitter.

$$\hat{x}_s = \sqrt{2}\hat{x}_{in} + \hat{x}_{id}^{(0)}, \quad (2.101)$$

which is a noisy version of the coherent state with twice the intensity. Then by splitting the signal output in two modes using a beamsplitter, as shown in Fig. 2.12, we obtain

$$\hat{x}_s = \hat{x}_{in} + \frac{1}{\sqrt{2}}\hat{x}_{id}^{(0)} + \frac{1}{\sqrt{2}}\hat{x}_v^{(0)}. \quad (2.102)$$

This is exactly a state with the same mean as the input state with one added unit of shot noise ($\gamma_{out} = 2\mathbb{I}$), which was previously shown to be the optimal cloning. Using equation (2.21) we see that the fidelity reads $F = 2/3$.

Classical Cloning We call classical cloning an inefficient way of making the copies which consists in optimally measuring the state, using heterodyne detection, and then re-preparing the copies. The fidelity that be obtain by classical cloning reads $F_C = 1/2$ which is lower than the one achieved using quantum operations ($F_Q = 2/3$).

Idler output The idler output,

$$\hat{x}_{id} = \hat{x} + \sqrt{2}\hat{x}_{id}^{(0)}, \quad (2.103)$$

$$\hat{p}_{id} = -\hat{p} + \sqrt{2}\hat{p}_{id}^{(0)}, \quad (2.104)$$

can be seen a noisy version of the phase conjugation operation $((\hat{x}, \hat{p}) \rightarrow (\hat{x}, -\hat{p}))$ applied to the input state [42]. One can also show that this phase conjugation operation being non-Hermitian (related to time-reversal) cannot be physically implemented. But as for cloning, one can make the phase conjugation feasible by allowing a noisy version of it. Strikingly the optimal operation can be obtained by using the output of the idler mode, reaches a fidelity of $F = 1/2$ as the idler output has an added noise of two shot noise units. Contrary to the cloning case, doing the operation quantumly does not gives better performance than doing it classically (measurement and re-preparation) as the optimal fidelity in both cases is $F = 1/2$.

Linear Optics Cloning Machine

A recent striking result was the discovery that the cloning machine can be implemented using just linear optics, homodyne measurement and controlled displacements [4], which is experimentally easier to implement than using OPA crystals and reaches a higher amplification gain. The idea is to replace the two-mode squeezer by the scheme presented in Fig. 2.13. A fraction of the input mode (\hat{x}_{in}) is reflected in a beamsplitter (T) and

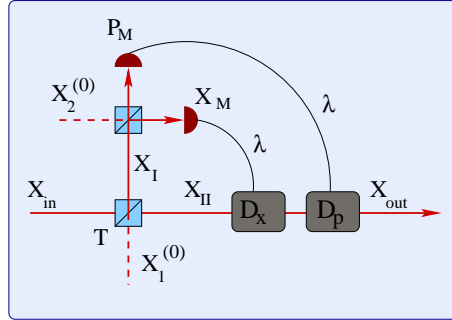


Figure 2.13: Fraction of the input mode (\hat{x}_{in}) is reflected in a beamsplitter (T) and subsequently measured with an heterodyne measurement. The transmitted mode (\hat{x}_{II}) is displaced depending on the measurement result (\hat{x}_M) and a the gain λ . By correctly choosing we obtain the transformation of a linear amplifier on the signal mode.

(T)

$$\hat{x}_I = \sqrt{T}\hat{x}_{in} + \sqrt{1-T}\hat{x}_1^{(0)}, \quad (2.105)$$

and subsequently measured with an heterodyne measurement

$$\hat{x}_M = \frac{1}{\sqrt{2}}(\hat{x}_I + \hat{x}_2^{(0)}) = \sqrt{\frac{T}{2}}\hat{x}_{in} + \sqrt{\frac{1-T}{2}}\hat{x}_1^{(0)} + \frac{1}{\sqrt{2}}\hat{x}_2^{(0)}. \quad (2.106)$$

The transmitted mode (\hat{x}_{II})

$$\hat{x}_{II} = \sqrt{1-T}\hat{x}_{in} - \sqrt{T}\hat{x}_1^{(0)} \quad (2.107)$$

is displaced depending on the measurement result (\hat{x}_M) and the gain λ ,

$$\hat{x}_{out} = \hat{x}_{II} + \lambda\hat{x}_M = \left[\sqrt{1-T} + \frac{\lambda\sqrt{T}}{\sqrt{2}}\right]\hat{x}_{in} + \left[\frac{\lambda\sqrt{1-T}}{\sqrt{2}} - \sqrt{T}\right]\hat{x}_1^{(0)} + \frac{\lambda}{\sqrt{2}}\hat{x}_2^{(0)}. \quad (2.108)$$

If we choose the gain such that $\lambda = \sqrt{2T}/\sqrt{1-T}$ the contribution of $\hat{x}_1^{(0)}$ is null, obtaining

$$\hat{x}_{out} = \frac{1}{\sqrt{1-T}}\hat{x}_{in} + \sqrt{\frac{T}{1-T}}\hat{x}_2^{(0)}, \quad (2.109)$$

which by setting $\sqrt{G} = 1/\sqrt{1-T}$ becomes exactly the operation of a linear amplifier on the signal mode.

Cloning and Measurement

The previous cloning machine of coherent states has as input a single copy and generates two noisy copies, it is then called a $1 \rightarrow 2$ cloning machine. One can then generalize the previous results to $N \rightarrow M$ cloning machines where we have N copies at the input and we output $M \geq N$ copies. The optimal fidelity reads,

$$F_{N \rightarrow M} = \frac{MN}{MN + M - N}, \quad (2.110)$$

as shown in [41, 42, 78]. It is interesting to point that in limit of infinite outputs ($M \rightarrow \infty$) the fidelity reads

$$F_{N \rightarrow \infty} = \frac{N}{N + 1}, \quad (2.111)$$

which is the limit that can be obtained with an optimal measurement [100, 103]. This shows that the measurement can be seen as a $N \rightarrow \infty$ cloning process.

Chapter 3

Generation of Arbitrary Single-Mode States

3.1 Introduction

In recent years, it has been widely recognized that nonclassical states of light represent a valuable resource for numerous applications ranging from ultra-high precision measurements [101, 102, 200, 201] to quantum lithography [23, 30] and quantum information processing [31]. It is often desirable to generate nonclassical states of traveling optical modes, as opposed to the cavity QED experiments where the generated state is confined in a cavity and can be probed only indirectly. Many ingenious schemes have been proposed and experimentally demonstrated to generate the single-photon states [130, 202] and various multiphoton entangled states such as the GHZ states [143, 144], cluster states [193], and the so-called NOON states [70, 79, 88, 125, 132].

Considerable attention has been also devoted to the preparation of arbitrary single-mode states [47, 54, 55, 191, 204] and, in particular, the Schrödinger cat-like superpositions of coherent states [53, 129] which can be useful for quantum information processing [111, 153]. The experimental generation of arbitrary superpositions of vacuum and single-photon states has been accomplished using parametric down-conversion with the input signal mode prepared in a coherent state [159], employing the quantum scissors scheme [12, 120, 147], or conditioning on homodyne measurements on one part of a non-local single photon in two spatial modes [11]. It is, however, very difficult to extend these experiments to superpositions involving two or more photons. The known schemes for conditional generation of arbitrary superpositions of Fock states, as the proposal of Dakna *et al.*, require single-photon sources and/or highly efficient detectors with single-photon resolution, which represents a formidable experimental challenge.

Dakna Proposal

Dakna *et al.* showed in [55] that any superposition of the first $N + 1$ Fock states of a single-mode of light,

$$|\psi\rangle = \sum_{n=0}^N c_n |n\rangle \quad (3.1)$$

can be engineered starting from vacuum by applying a sequence of displacements and single-photon additions, as shown in Fig. 3.1. Remembering that the definition of an

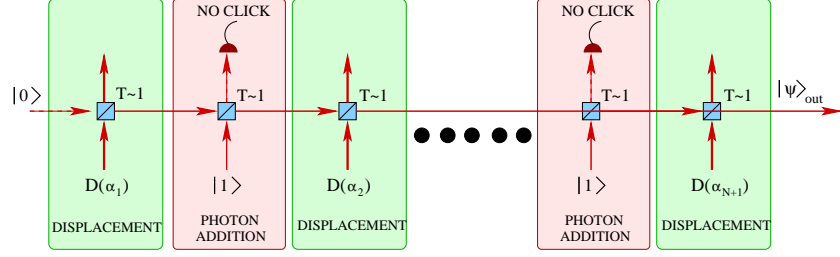


Figure 3.1: Dakna *et al.* proposal for experimentally generating a superposition of Fock states is based on a sequence of displacements interspersed with a conditional photon subtraction. A successful state preparation is heralded by no click of the photodetectors (APD).

arbitrary single-mode state (3.1) can be rewritten as

$$\prod_{i=1}^N (\hat{a}^\dagger - \beta_i^*) |0\rangle, \quad (3.2)$$

where β_i^* are the N (complex) roots of the characteristic polynomial $\sum_{k=0}^N c_k \hat{a}^k / \sqrt{n!}$. Using the relation

$$\hat{a}^\dagger - \beta^* = D(\beta) \hat{a}^\dagger D^\dagger(\beta), \quad (3.3)$$

we find that

$$|\psi\rangle = D(\beta_N) \hat{a}^\dagger D^\dagger(\beta_N) D(\beta_{N-1}) \hat{a}^\dagger D^\dagger(\beta_{N-1}) \times \dots \times D(\beta_1) \hat{a}^\dagger D^\dagger(\beta_1) |0\rangle. \quad (3.4)$$

Because $D^\dagger(\beta_N) D(\beta_{N-1}) = D(\beta_{N-1} - \beta_N)$, we can obtain any quantum state (3.1) from the vacuum by a succession of alternate state displacement and single-photon addition.

The two major drawbacks of this proposal are: Firstly, the need of single-photon sources on demand, which is experimentally very challenging, in order to implement the photon addition step of the protocol; Secondly, because the successful state preparation is heralded by a no click of the photodetectors, the scheme is extremely sensitive to the inefficiency of the photodetectors.

New Proposal

In this chapter, we propose a novel state preparation scheme inspired by the proposal of Dakna *et al.* [55], which does not require single-photon sources and can operate with high fidelity even with low-efficiency detectors that only distinguish the presence or absence of photons. Our crucial observation is that if the initial state is a squeezed vacuum, then the single-photon addition can be replaced with single-photon subtraction [139, 49, 115], as shown in Fig. 3.2, which is much more practicable. Indeed, a single-photon subtraction can be achieved by diverting a tiny fraction of the beam with a beam splitter towards a photodetector, so that a click means that a photon has been subtracted from the beam (this process becomes exact for a transmittance tending to one). In fact, the single-photon subtraction from a squeezed vacuum has already been experimentally demonstrated [197], which provides a strong evidence for the practical feasibility of our scheme. We note that the photon subtraction is an extremely useful tool which allows one to generate states suitable for the tests of Bell inequality violation with balanced homodyning [85, 135]. It can also be used to improve the performance

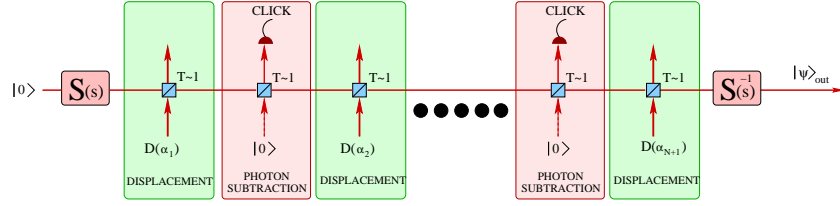


Figure 3.2: Our proposal for experimentally generating a superposition of Fock states is based on a sequence of displacements interspersed with a conditional photon subtraction applied to a squeezed vacuum state $S(s)|0\rangle$. A successful state preparation is heralded by a click of all photodetectors (APD).

of dense coding [118], and forms a crucial element of the entanglement distillation schemes for continuous variables [37]. However, the counterpart of using photon subtraction in our preparation scheme lies in that a final anti-squeezing operation needs to be performed. The implementation of this operation is technically more involved than the initial squeezed vacuum preparation, although it has already been experimentally demonstrated [17, 18, 126, 151]. In addition, a new method based on homodyne detection followed by a feed-forward displacement has been proposed recently [77].

The present chapter is organized as follows. In Section 3.2, we explain the mechanism of state generation on the simplest non-trivial example of a superposition of vacuum and single-photon states. Our setup then consists of two displacements, one conditional photon subtraction, and two squeezers (squeezing conjugate quadratures). We present the details of the calculation of the Wigner function of the generated state for a realistic setup involving imperfect photon subtraction (obtained with imperfect detectors and beam splitters with a non-unity transmittance). In order to evaluate the performance of the scheme, we investigate the achieved fidelity and the preparation probability for various target states. We also discuss the feasibility of the final squeezing operation. In Section 3.3, we extend the scheme to the generation of an arbitrary single-mode state and show how to calculate the displacements that need to be applied during the state preparation. As an illustration, we consider the generation of several states which are superpositions of vacuum, single-photon, and two-photon Fock states. In Section 3.4, we propose an iterative state generation scheme that uses a quantum memory in order to reduce very significantly the total number of required operations.

3.2 Generation of a Superposition of $|0\rangle$ and $|1\rangle$

In this Section, we introduce our setup for the generation of an arbitrary superposition of vacuum and single-photon state, which consists of two squeezers and two displacements with a photon subtraction in between, as schematically sketched in Fig. 3.3. This setup represents a basic building block of our universal scheme: as shown in Section 3.3, any superposition of the first $N + 1$ Fock states can be generated from a single-mode squeezed vacuum by a displacement followed by a sequence of N photon subtractions and displacements, completed by a final anti-squeezing operation.

Pure-State Description

We first provide a simplified pure-state description of the setup, assuming perfect detectors with single-photon resolution, which will give us an insight into the mechanism of state generation. We will show that, conditionally on observing a click of the photode-

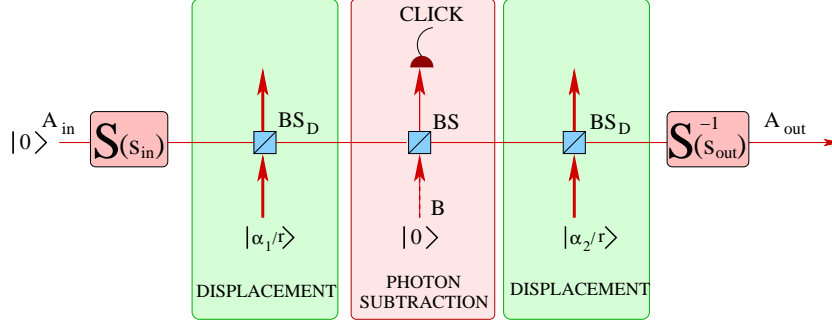


Figure 3.3: Proposed experimental setup for generating $|\psi\rangle = c_0|0\rangle + c_1|1\rangle$. An optical parametric amplifier generates a single-mode squeezed vacuum state (of squeezing parameter s_{in}), which then propagates through three highly unbalanced beam splitters (BS_D , BS , and BS_D) in order to realize a sequence of two displacements interspersed with one conditional photon subtraction. Finally, an anti-squeezing $S^\dagger(s_{\text{out}})$ operation is applied, resulting in the output mode A_{out} . Successful state preparation is heralded by a click of the photodetectors (APD).

tector APD, the setup produces a superposition of vacuum and single-photon states,

$$|\psi\rangle_{\text{target}} = c_0|0\rangle + c_1|1\rangle. \quad (3.5)$$

Our state engineering procedure starts with a single-mode squeezed vacuum state, which is generated in an optical parametric amplifier,

$$S(s_{\text{in}})|0\rangle = \frac{1}{\sqrt{\cosh(s_{\text{in}})}} \sum_{n=0}^{\infty} \frac{\sqrt{(2n)!}}{2^n n!} [\tanh(s_{\text{in}})]^n |2n\rangle, \quad (3.6)$$

where $S(s) = \exp[s(a^{\dagger 2} - a^2)/2]$ denotes the squeezing operator with a (a^\dagger) being the annihilation (creation) operator, and s_{in} denotes the initial squeezing constant. The single-mode squeezed vacuum passes through three highly transmitting beam splitters, which realize a sequence consisting of a displacement followed by a single-photon subtraction and another displacement. The state is displaced by combining it on a highly unbalanced beam splitter BS_D with transmittance $T_D > 99\%$ with a strong coherent state $|\alpha/r_D\rangle$, where $r_D = \sqrt{R_D}$ and $R_D = 1 - T_D$ is the reflectance of BS_D [145]. In the limit $T_D \rightarrow 1$, the output beam is displaced by the amount α . This method has been used, e.g., in the continuous-variable quantum teleportation experiment [83]. For the sake of simplicity, we shall assume that $T_D = 1$ and the displacement operation is exact. The conditional single-photon subtraction requires a highly unbalanced beam splitter BS with transmittance T , followed by a photodetector PD placed on the auxiliary output port. A successful photon subtraction is heralded by a click of the detector. In the limit $T \rightarrow 1$, the most probable event leading to a click of the detector is that exactly a single photon has been reflected from the beam splitter. The probability of removing two or more photons is smaller by a factor of $1 - T$ and becomes totally negligible in the limit $T \rightarrow 1$. The conditional single-photon subtraction can be described by the non-unitary operator

$$\hat{X} = t^{\hat{n}} r \hat{a}, \quad (3.7)$$

where $\hat{n} = \hat{a}^\dagger \hat{a}$ is the photon-number operator, while $t = \sqrt{T}$ and $r = \sqrt{1 - T}$ denote the amplitude transmittance and reflectance of BS , respectively.

The input-output transformation corresponding to the sequence of operations in Fig. 3.2 reads

$$|\psi\rangle_{\text{out}} = S^\dagger(s_{\text{out}}) D(\alpha_2) \hat{X} D(\alpha_1) S(s_{\text{in}}) |0\rangle, \quad (3.8)$$

where $D(\alpha) = \exp(\alpha\hat{a}^\dagger - \alpha^*\hat{a})$ is the displacement operator. We will show later on how the displacements α_1 and α_2 depend on the target state (3.5), as well as how the initial squeezing value s_{in} depends on the output squeezing s_{out} for a given transmittance $T < 1$.

In order to give an intuitive description of the technique let us assume first that $T = 1$, $\alpha_1 = -\alpha_2 = \alpha$, $s_{\text{out}} = s_{\text{in}} = s$, and replace \hat{X} by \hat{a} . Then, using $D(\alpha)^\dagger \hat{a} D(\alpha) = \hat{a} + \alpha$, the conditionally generated state can be written as

$$|\psi\rangle_{\text{out}} = S^\dagger(s) (\hat{a} + \alpha) S(s) |0\rangle. \quad (3.9)$$

Taking into account that \hat{a} and \hat{a}^\dagger transform under the squeezing operation according to

$$\begin{aligned} S^\dagger(s) \hat{a} S(s) &= \hat{a} \cosh(s) + \hat{a}^\dagger \sinh(s), \\ S^\dagger(s) \hat{a}^\dagger S(s) &= \hat{a}^\dagger \cosh(s) + \hat{a} \sinh(s), \end{aligned} \quad (3.10)$$

we obtain

$$\begin{aligned} |\psi\rangle_{\text{out}} &= [\hat{a} \cosh(s) + \hat{a}^\dagger \sinh(s) + \alpha] |0\rangle \\ &= \sinh(s) |1\rangle + \alpha |0\rangle. \end{aligned} \quad (3.11)$$

We can see that by setting $\alpha = (c_0/c_1) \sinh(s)$, we obtain the target state (3.5). This simple analysis illustrates the principle of the scheme shown in Fig. 3.2.

However, the limit $T = 1$ is unphysical, because the probability of successful state generation vanishes when $T \rightarrow 1$. Let us now estimate the realistic values of the displacements $\alpha_{1,2}$ taking into account that $T < 1$. In order to simplify the expression (3.8), we first rewrite all displacement operators in a normally-ordered form, $D(\alpha) = e^{-|\alpha|^2/2} e^{\alpha\hat{a}^\dagger} e^{-\alpha^*\hat{a}}$, and we obtain

$$|\psi\rangle_{\text{out}} \propto S^\dagger(s_{\text{out}}) e^{\alpha_2 \hat{a}^\dagger} e^{-\alpha_2^* \hat{a}} t^{\hat{n}} \hat{a} e^{\alpha_1 \hat{a}^\dagger} e^{-\alpha_1^* \hat{a}} S(s_{\text{in}}) |0\rangle. \quad (3.12)$$

Next, we propagate the operator $t^{\hat{n}}$ to the right by using the relations (see Appendix C)

$$t^{\hat{n}} e^{\alpha^* \hat{a}} = e^{\alpha^* \hat{a}/t} t^{\hat{n}}, \quad t^{\hat{n}} e^{\alpha \hat{a}^\dagger} = e^{t\alpha \hat{a}^\dagger} t^{\hat{n}}, \quad t^{\hat{n}} \hat{a} = \hat{a} t^{\hat{n}-1}. \quad (3.13)$$

After these algebraic manipulations we obtain

$$|\psi\rangle_{\text{out}} \propto S^\dagger(s_{\text{out}}) e^{\alpha_2 \hat{a}^\dagger} \hat{a} e^{t\alpha_1 \hat{a}^\dagger} e^{-[\alpha_2^* + \alpha_1^*/t] \hat{a}} t^{\hat{n}} S(s_{\text{in}}) |0\rangle. \quad (3.14)$$

Note that we have also moved to the right the operator $e^{-\alpha_2^* \hat{a}}$, used the fact that $e^{\alpha \hat{a}^\dagger} e^{\beta^* \hat{a}} = e^{-\alpha\beta^*} e^{\beta^* \hat{a}} e^{\alpha \hat{a}^\dagger}$ and dropped some t and r as we are only interested in the target state and not the success probability.

The combined action of the operators $t^{\hat{n}} S(s_{\text{in}})$ on vacuum produces a single-mode squeezed vacuum state just as without applying $t^{\hat{n}}$ but with a lower squeezing constant s satisfying

$$\tanh(s) = t^2 \tanh(s_{\text{in}}), \quad (3.15)$$

that is

$$t^{\hat{n}} S(s_{\text{in}}) |0\rangle \propto S(s) |0\rangle. \quad (3.16)$$

Finally, we move the operator $e^{\alpha_2 \hat{a}^\dagger}$ to the right, using the formula $e^{\alpha_2 \hat{a}^\dagger} \hat{a} = (\hat{a} - \alpha_2) e^{\alpha_2 \hat{a}^\dagger}$, which results in

$$|\psi\rangle_{\text{out}} \propto S^\dagger(s_{\text{out}}) (\hat{a} - \alpha_2) e^{\delta \hat{a}^\dagger} e^{-\gamma^* \hat{a}} S(s) |0\rangle, \quad (3.17)$$

where $\delta = \alpha_2 + t\alpha_1$ and $\gamma = \alpha_2 + \alpha_1/t$. With the help of Eq. (3.10), we can write, using $S(s)S(s)^\dagger = \mathbb{I}$,

$$e^{\delta\hat{a}^\dagger}e^{-\gamma^*\hat{a}}S(s)|0\rangle \propto S(s)S(s)^\dagger e^{\delta\hat{a}^\dagger}e^{-\gamma^*\hat{a}}S(s)|0\rangle \quad (3.18)$$

$$\propto S(s)e^{[\delta(\cosh(s)\hat{a}^\dagger + \sinh(s)\hat{a}) - \gamma^*(\cosh(s)\hat{a} + \sinh(s)\hat{a}^\dagger)]}|0\rangle, \quad (3.19)$$

$$\propto S(s)e^{[\delta\cosh(s) - \gamma^*\sinh(s)]\hat{a}^\dagger}|0\rangle, \quad (3.20)$$

which is a state with a generally non-zero coherent displacement. This displacement can be set to zero if α_1 and α_2 satisfy

$$(\alpha_2 + t\alpha_1)\cosh(s) = (\alpha_2^* + \alpha_1^*/t)\sinh(s), \quad (3.21)$$

in which case the output state reads

$$|\psi\rangle_{\text{out}} \propto S^\dagger(s_{\text{out}})(\alpha_2 - \alpha_2^*)S(s)|0\rangle \quad (3.22)$$

Finally, if we choose $s_{\text{out}} = s$, using Eq. (3.10) we obtain,

$$|\psi\rangle_{\text{out}} \propto \sinh(s)|1\rangle - \alpha_2|0\rangle. \quad (3.23)$$

Thus, the desired superposition of the first two Fock states (3.5) can be obtained by choosing

$$\alpha_2 = -\frac{c_0}{c_1}\sinh(s), \quad (3.24)$$

$$\alpha_1 = t\frac{[\tanh^2(s) - t^2]\alpha_2 + (t^2 - 1)\tanh(s)\alpha_2^*}{t^4 - \tanh^2(s)}, \quad (3.25)$$

where the displacement α_1 has been determined from the condition (3.21) by subtracting t times (3.21) and the conjugate of (3.21) divided by t (in order to suppress the term α_1^*), followed by some rearranging. Note that we may assume that the coefficient c_1 of the Fock state $|1\rangle$ is non-zero; otherwise, no photon subtraction is needed to generate the target state.

Final Anti-Squeezing Operation

In order to obtain a superposition of Fock states at the output, we need to apply the final anti-squeezing operation $S^\dagger(s_{\text{out}})$, which squeezes a quadrature conjugate to that squeezed by the first squeezer $S(s_{\text{in}})$. This operation can be implemented by injecting the signal beam into a nonlinear medium that is strongly pumped by a laser, as demonstrated in [17, 18, 126, 151]. A difficulty of this method lies in that a good spatio-temporal overlap between the signal and the pump beams must be achieved. However, a recently proposed alternative method can be used to avoid this problem. Here, an auxiliary mode that is prepared in a squeezed vacuum state is combined with the signal beam at a beam splitter. The auxiliary mode is then measured with a homodyne detector and the appropriate quadrature of the signal beam is displaced proportionally to the measurement outcome [77]. The great advantage of this latter approach is that it only requires the interference between two beams at a beam splitter, which is much easier to achieve than the direct phase-sensitive de-amplification of the signal in a nonlinear medium. A very similar scheme has been in fact successfully demonstrated in the recent experiment of continuous variable quantum erasing [3].

Note that if we remove the last squeezing operation $S^\dagger(s_{\text{out}})$, we obtain a simpler optical setup which produces a squeezed superposition of Fock states $S(s_{\text{out}})[c_0|0\rangle + c_1|1\rangle]$. In many cases, however, this squeezing may not be an obstacle or may even represent an advantage. For example, the generation of Schrödinger cat states $|\alpha\rangle - |-\alpha\rangle$ can, for small $|\alpha|$, be very well approximated by a squeezed single-photon state $S(s)|1\rangle$ [115, 129].

Realistic Model

We shall now present a more realistic description of the proposed scheme taking into account that the photodetectors exhibit only single-photon sensitivity, but cannot resolve the number of photons in the mode, and have a detection efficiency $\eta < 1$. Such detectors have two outcomes, either a click or a no-click. We model this detector as a sequence consisting of a beam splitter with transmittance η followed by an idealized detector which performs a projection onto the vacuum or the rest of the Hilbert space, $\Pi_0 = |0\rangle\langle 0|$ (no click), $\Pi_1 = \mathbb{I} - |0\rangle\langle 0|$ (a click).

Similarly as in Ref. [85], it is convenient to work in the phase-space representation and consider the transformation of Wigner functions. The setup in Fig. 1 involves two modes, the principal mode A and an auxiliary mode B. Initially, the mode A is in a squeezed vacuum state and the covariance matrix is diagonal, $\gamma_A = \text{diag}(e^{-2s_{\text{in}}}, e^{2s_{\text{in}}})$. The Gaussian Wigner function of the initial state of mode A after the first displacement thus reads

$$W_G(r_A; \Gamma_A, d_A) = \frac{\sqrt{\det \Gamma_A}}{\pi} e^{-(r_A - z_1)^T \Gamma_A (r_A - z_1)}, \quad (3.26)$$

where $r_A = (x_A, p_A)^T$ is the vector of quadratures of mode A and $z_1 \equiv \sqrt{2}(\Re \alpha_1, \Im \alpha_1)^T$ is the displacement. The matrix Γ_A is the inverse of the covariance matrix γ_A .

In a second step, the modes A and B are mixed on an unbalanced beam splitter BS and then mode B subsequently passes through a (virtual) beam splitter of transmittance η which models the imperfect detection with efficiency η . This transformation is a Gaussian completely positive (CP) map \mathcal{M} , and the resulting state of modes A and B is still a Gaussian state with the Wigner function

$$W_{AB}(r_{AB}) = \frac{\sqrt{\det \Gamma_{AB}}}{\pi^2} e^{-(r_{AB} - d_{AB})^T \Gamma_{AB} (r_{AB} - d_{AB})}, \quad (3.27)$$

where $r_{AB} = (r_A, r_B)^T$. The vector of the first moments $d_{AB} = (d_A, d_B)^T$ and the covariance matrix $\gamma_{AB} = \Gamma_{AB}^{-1}$ can be expressed in terms of the initial parameters of mode A before the mixing on an unbalanced beam splitter (BS) (i.e., z_1 and γ_A) as follows:

$$\begin{aligned} d_{AB} &\equiv \begin{pmatrix} d_A \\ d_B \end{pmatrix} = S \begin{pmatrix} z_1 \\ 0 \end{pmatrix}, \\ \gamma_{AB} &= S(\gamma_A \oplus I_B)S^T + G, \end{aligned} \quad (3.28)$$

where $S = S_\eta S_{BS}$, $S_\eta = I_A \oplus \sqrt{\eta} I_B$ and $G = 0_A \oplus (1 - \eta) I_B$ model the inefficient photodetector, and S_{BS} is a symplectic matrix which describes the coupling of the modes A and B on an unbalanced beam splitter (BS),

$$S_{BS} = \begin{bmatrix} t & 0 & r & 0 \\ 0 & t & 0 & r \\ -r & 0 & t & 0 \\ 0 & -r & 0 & t \end{bmatrix}. \quad (3.29)$$

After the photon subtraction, the density matrix $\rho_{A,\text{out}}$ of mode A conditioned on a click of the photodetector PD measuring the auxiliary mode B becomes

$$\rho_{A,\text{out}} = \text{Tr}_B[\rho_{AB}(\mathbb{I}_A \otimes \Pi_{1,B})], \quad (3.30)$$

where Tr_B denotes a partial trace over mode B, and ρ_{AB} is the two-mode density matrix of the Gaussian state characterized by the Wigner function (3.27). Then, after the second displacement of $z_2 \equiv \sqrt{2}(\Re \alpha_2, \Im \alpha_2)^T$, the Wigner function of mode A can be written as a linear combination of two Gaussian functions (3.26), namely

$$W(r) P = C_1 W_G(r; \Gamma_1, d_1) + C_2 W_G(r; \Gamma_2, d_2), \quad (3.31)$$

where P is the probability of successful generation of the target state. The expression (3.31) can be derived by rewriting Eq. (3.30) in the Wigner representation. One uses the fact that the Wigner function of the POVM element $\Pi_{1,B}$ is a difference of two Gaussian functions,

$$W_{\Pi_1}(r) = \frac{1}{2\pi} - \frac{1}{\pi} e^{-x^2 - p^2}, \quad (3.32)$$

and that the trace of the product of two operators can be evaluated by integrating the product of their Wigner representations over the phase space.

To define the matrices and vectors appearing in Eq. (3.31), we first divide the matrix $\Gamma_{AB} = \gamma_{AB}^{-1}$ into four sub-matrices with respect to the $A|B$ splitting,

$$\Gamma_{AB} = \begin{bmatrix} \Upsilon_A & \sigma \\ \sigma^T & \Upsilon_B \end{bmatrix}, \quad (3.33)$$

and we carry some lengthly but simple algebra, as shown in Appendix D. As detailed in Appendix D, the correlation matrix Γ_1 and the displacement d_1 appearing in the first term on the right-hand side of Eq. (3.31) are given by

$$\begin{aligned} \Gamma_1 &= \Upsilon_A - \sigma \Upsilon_B^{-1} \sigma^T, \\ d_1 &= d_A + z_2, \\ C_1 &= 1. \end{aligned} \quad (3.34)$$

Similarly, the formulas for the parameters of the second term read

$$\begin{aligned} \Gamma_2 &= \Upsilon_A - \sigma \tilde{\Upsilon}_B^{-1} \sigma^T, \\ d_2 &= d_A + \Gamma_2^{-1} \sigma \tilde{\Upsilon}_B^{-1} d_B + z_2, \\ C_2 &= -2 \sqrt{\frac{\det(\Gamma_{AB})}{\det(\Gamma_2) \det(\tilde{\Upsilon}_B)}} \exp[-d_B^T M d_B], \end{aligned} \quad (3.35)$$

where $\tilde{\Upsilon}_B = \Upsilon_B + I$ and

$$M = \Upsilon_B \tilde{\Upsilon}_B^{-1} - \tilde{\Upsilon}_B^{-1} \sigma^T \Gamma_2^{-1} \sigma \tilde{\Upsilon}_B^{-1}. \quad (3.36)$$

The final squeezing operation $S^\dagger(s_{\text{out}})$, described by the symplectic matrix

$$S_s = \begin{bmatrix} e^{s_{\text{out}}} & 0 \\ 0 & e^{-s_{\text{out}}} \end{bmatrix}, \quad (3.37)$$

is applied to mode A after the last displacement. The resulting Wigner function of the output mode A_{out} can be written as

$$W_{\text{out}}(r) P = C_1 W_G(r; \Gamma'_1, d'_1) + C_2 W_G(r; \Gamma'_2, d'_2), \quad (3.38)$$

where the inverse covariance matrix $\Gamma'_{1,2}$ and the displacement $d_{1,2}$ appearing in the right-hand side of Eq. (3.38) are given by

$$\begin{aligned} \Gamma'_{1,2} &= S_s^{-1} \Gamma_{1,2} S_s^{-1}, \\ d'_{1,2} &= S_s d_{1,2}. \end{aligned} \quad (3.39)$$

Since all the Wigner functions in Eq. (3.31) or (3.38) are normalized, the probability of a successful state generation can be calculated simply as the sum $P = C_1 + C_2$.

Examples

In order to illustrate our method, let us consider the preparation of the following four superpositions of the Fock states $|0\rangle$ and $|1\rangle$,

$$|\psi_1\rangle = |1\rangle, \quad (3.40)$$

$$|\psi_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \quad (3.41)$$

$$|\psi_3\rangle = \frac{1}{\sqrt{10}}(3|0\rangle + |1\rangle), \quad (3.42)$$

$$|\psi_4\rangle = \frac{1}{\sqrt{10}}(|0\rangle + 3|1\rangle). \quad (3.43)$$

In the following we compare the fidelity of the state obtained by the realistic implementation ($W_r(r)$) to the ideal pure state ($W_i(r)$),

$$F = 2\pi \int dr W_r(r) W_i(r), \quad (3.44)$$

where the pure state Wigner function is calculated as shown in Appendix E.

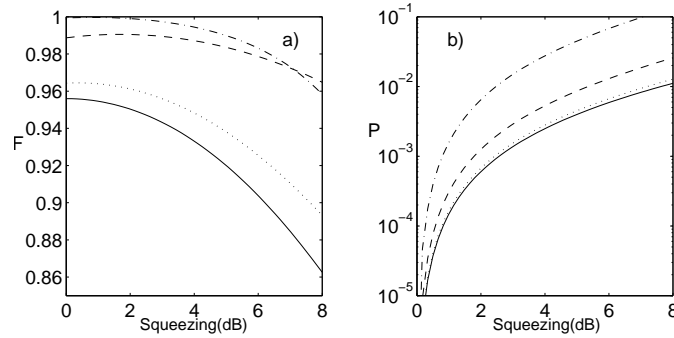


Figure 3.4: (a) Fidelity between the generated state and the target state and (b) probability of successful generation as a function of the squeezing s_{in} for the four target states (3.40) (solid line), (3.41) (dashed line), (3.42) (dot-dashed line) and (3.43) (dotted line), with $T = 0.95$ and $\eta = 0.25$.

The fidelity of the generated state for the target states (3.40)–(3.43) is plotted in Fig. 3.4(a) as a function of the initial squeezing. We can see that the conditionally prepared states are close to the desired states and their optimum fidelities are reached for a low initial squeezing (below 2 dB), which is experimentally accessible. Although it is hardly visible in Fig 3.4(a), there is typically a non-zero optimal value of the initial squeezing, giving the highest fidelity. As shown in Fig. 3.4(b) the increase of the initial squeezing improves the probability of successful generation of the target state. A comparison of Fig. 3.4(a) with Fig. 3.4(b) reveals a clear trade-off between the achievable fidelity and the preparation probability.

Figure 3.5(a) shows the dependence of the fidelity on the beam splitter transmittance T , considering the optimal input squeezing for each of the states. (Note that for state (3.40), we could not find numerically the optimum squeezing, so we arbitrarily chose $s_{\text{in}} = 0.50$ dB as an optimal value in order to keep a reasonable generation probability.) We see that as T approaches unity, the fidelity gets arbitrarily close to unity, while the probability of successful state generation decreases as $P \propto (1 - T)\eta$, as shown in Fig. 3.5(b). The value $T = 0.95$ used in Fig. 3.4 seems to be a reasonable

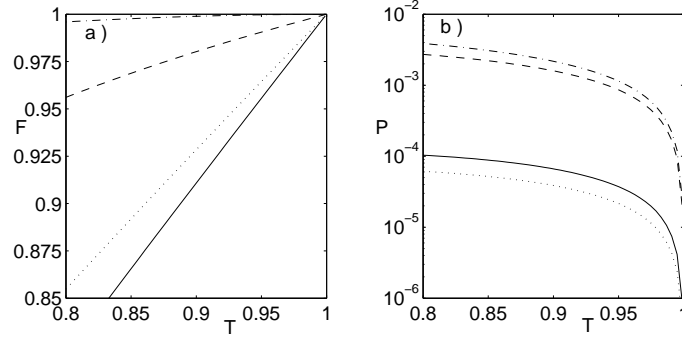


Figure 3.5: (a) Fidelity between the generated state and the target state and (b) probability of successful generation as a function of T for the four target states (3.40)–(3.43). The curves are plotted considering the optimal squeezing s_{in} for each state, namely 0.50 dB for state (3.40) (solid line), 1.66 dB for state (3.41) (dashed line), 0.85 dB for state (3.42) (dot-dashed line), and 0.36 dB for state (3.43) (dotted line). The curves are plotted for $\eta = 0.25$.

compromise between the success rate ($P \approx 10^{-3}$ or $P \approx 10^{-4}$ depending on the target state) and the fidelity $F > 0.95\%$.

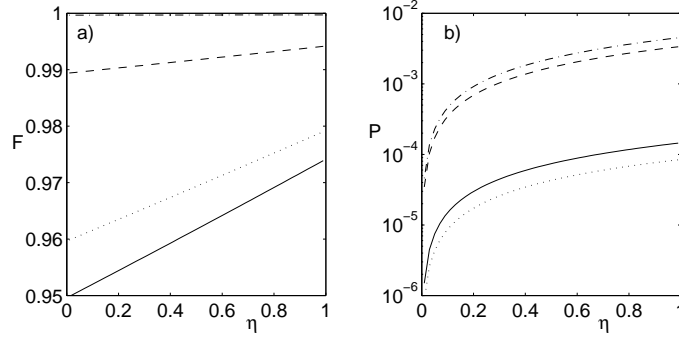


Figure 3.6: (a) Fidelity between the generated state and the target state and (b) probability of successful generation as a function of η for the four target states (3.40)–(3.43). The curves are plotted considering the optimal squeezing s_{in} for each state, namely 0.50 dB for state (3.40) (solid line), 1.66 dB for state (3.41) (dashed line), 0.85 dB for state (3.42) (dot-dashed line), and 0.36 dB for state (3.43) (dotted line). The curves are plotted for $T = 0.95$.

We also have studied the dependence of the fidelity on the detection efficiency η . The numerical results are shown in Fig. 3.6(a), where we can see that the scheme is very robust in the sense that the fidelity almost does not depend on η . Fidelities above 95% could be reached even with η of the order of a few percents if T is high enough. This is in agreement with the findings of Ref. [85]. However, a low η reduces the preparation probability, as shown in Fig. 3.6(b).

3.3 Arbitrary Single-Mode State

In the preceding section, we have demonstrated that the combination of two displacements and a photon subtraction allows us to build any superposition of $|0\rangle$ and $|1\rangle$ states. In this section, we shall generalize this procedure to any superposition of the first $N + 1$ Fock states,

$$|\psi\rangle_{\text{target}} = \sum_{n=0}^N c_n |n\rangle, \quad (3.45)$$

and show that it can be prepared from a squeezed vacuum state by applying a sequence of $N + 1$ displacements interspersed with N photon subtractions, and a final anti-squeezing operation as shown in Fig. 3.7.

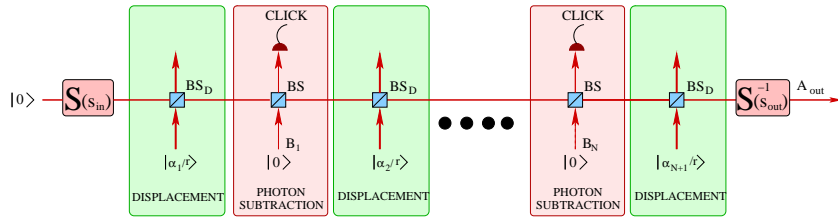


Figure 3.7: Proposed experimental setup. An optical parametric amplifier generates a single-mode squeezed vacuum state $S(s_{\text{in}})|0\rangle$, which then propagates through $2N + 1$ highly unbalanced beam splitters BS_D and BS , which realize a sequence of $N + 1$ displacements interspersed with N conditional photon subtractions. A second squeezer is used to apply the final anti-squeezing operation $S^\dagger(s_{\text{out}})$. Successful state preparation is heralded by clicks of all N photodetectors PD_k .

Pure-State Description

As in the preceding section, we first provide a simplified pure-state description of the setup, assuming perfect detectors with single-photon resolution. This will allow us to determine the dependence of the coherent displacements α_j on the target state (3.45). Generalizing the procedure presented in the preceding section, the input-output transformation corresponding to the sequence of operations in Fig. 3.7 reads

$$|\psi\rangle_{\text{out}} = S^\dagger(s_{\text{out}}) D(\alpha_{N+1}) \hat{X} D(\alpha_N) \hat{X} \dots D(\alpha_2) \hat{X} D(\alpha_1) S(s_{\text{in}}) |0\rangle. \quad (3.46)$$

In order to simplify this expression, we first rewrite all displacement operators in a normally-ordered form and then move all the operators $t^{\hat{n}}$ to the right using the relations (3.13). This results in the substitution $\alpha_j \rightarrow \alpha_j t^{N+1-j}$ and $\alpha_j^* \rightarrow \alpha_j^* t^{j-N-1}$ in the exponents. Next, we propagate all the exponential operators $e^{-t^{j-N-1} \alpha_j^* \hat{a}}$ to the right,

$$|\psi\rangle_{\text{out}} \propto S^\dagger(s_{\text{out}}) e^{\alpha_{N+1} \hat{a}^\dagger} \hat{a} e^{t \alpha_N \hat{a}^\dagger} \hat{a} \dots \hat{a} e^{t^N \alpha_1 \hat{a}^\dagger} e^{-\gamma^* \hat{a}} t^{N \hat{n}} S(s_{\text{in}}) |0\rangle, \quad (3.47)$$

where $\gamma = \sum_{j=1}^{N+1} \alpha_j t^{j-N-1}$. The combined action of the operators $t^{N \hat{n}} S(s_{\text{in}})$ on the vacuum produces a single-mode squeezed vacuum state, $t^{N \hat{n}} S(s_{\text{in}}) |0\rangle \propto S(s) |0\rangle$, where

$\tanh(s) = t^{2N} \tanh(s_{\text{in}})$. After some algebraic manipulations and taking $s_{\text{out}} = s$, we get

$$|\psi\rangle_{\text{out}} \propto \prod_{j=1}^N (\hat{a} \cosh(s) + \hat{a}^\dagger \sinh(s) - \beta_j) |0\rangle, \quad (3.48)$$

where

$$\beta_j = \sum_{k=j+1}^{N+1} \alpha_k t^{N+1-k}. \quad (3.49)$$

Formula (3.48) is valid provided that the overall displacement is zero, corresponding to the constraint

$$\cosh(s) \sum_{j=1}^{N+1} \alpha_j t^{N+1-j} = \sinh(s) \sum_{j=1}^{N+1} \alpha_j^* t^{j-N-1}, \quad (3.50)$$

which generalizes condition (3.21).

We now prove that an arbitrary superposition of the first $N+1$ Fock states $\sum_{n=0}^N c_n |n\rangle$ can be expressed as $\prod_{j=1}^N (A - \beta_j) |0\rangle \equiv \sum_{k=0}^N h_k A^k |0\rangle$, where $A = \hat{a} \cosh(s) + \hat{a}^\dagger \sinh(s)$ and h_k are the coefficients of the characteristic polynomial whose roots are β_j . From the condition

$$\sum_{k=0}^N h_k A^k |0\rangle = \sum_{n=0}^N c_n |n\rangle, \quad (3.51)$$

we can immediately determine the coefficients h_N and h_{N-1} . This is because only the term A^N gives rise to $\hat{a}^{\dagger N}$ and, similarly, only the expansion of A^{N-1} contains $\hat{a}^{\dagger N-1}$. We thus get

$$h_N = \frac{c_N \sinh^{-N}(s)}{\sqrt{N!}}, \quad h_{N-1} = \frac{c_{N-1} \sinh^{1-N}(s)}{\sqrt{(N-1)!}}. \quad (3.52)$$

Once we know h_N and h_{N-1} , we insert them back in Eq. (3.51), and, from $\sum_{k=0}^{N-2} h_k A^k |0\rangle = \sum_{n=0}^N c_n |n\rangle - (h_{N-1} A^{N-1} + h_N A^N) |0\rangle$, we determine h_{N-2} and h_{N-3} . By repeating this procedure, we can find all coefficients h_j . This proves that the condition (3.51) can always be met for any nonzero squeezing, hence our method is indeed universal and allows us to generate *arbitrary* superpositions. After finding the h_j 's, the coefficients β_j 's are calculated as the roots of the characteristic polynomial $\sum_{k=0}^N h_k \beta^k$, and, finally, the $N+1$ displacements α_j 's are determined by solving the system of $N+1$ linear equations (3.49) and (3.50).

Realistic Model

We shall now present a more realistic description of the proposed scheme, which takes into account realistic photodetectors. After the k -th photon subtraction and the $k+1$ -th displacement, the density matrix $\rho_{k,A}$ of mode A conditioned on a click of the photodetector measuring the auxiliary mode B_k is related to $\rho_{k-1,A}$ as follows,

$$\rho_{k,A} = D_{k+1} \text{Tr}_B [\mathcal{M}(\rho_{k-1,A} \otimes |0\rangle_B \langle 0|) (\mathbb{I}_A \otimes \Pi_{1,B})] D_{k+1}^\dagger, \quad (3.53)$$

where $\rho_{0,A} = S(s_{\text{in}}) |0\rangle \langle 0| S^\dagger(s_{\text{in}})$, $D_{k+1} = D(\alpha_{k+1})$ is a displacement operator and \mathcal{M} denotes the Gaussian CP map (3.28) that describes the mixing of the modes A and B on BS and accounts for imperfect detection. Since each step (3.53) gives rise to a

linear combination of two Gaussian states from a Gaussian state, the Wigner function of the state $\rho_{k,A}$ can be written as a linear combination of 2^k Gaussian functions,

$$W_k(r)P_k = \sum_{j=1}^{2^k} C_{j,k} W_G(r; \Gamma_{j,k}, d_{j,k}), \quad (3.54)$$

where P_k is the probability of success of the first k photon subtractions. The correlation matrices $\Gamma_{j,k}$ and displacements $d_{j,k}$ after k photon subtractions and $k+1$ displacements can be expressed in terms of $\Gamma_{j,k-1}$ and $d_{j,k-1}$.

Similarly as in Section 3.2, we first define the real displacement vector $z_k \equiv \sqrt{2}(\Re\alpha_k, \Im\alpha_k)^T$ and the two-mode covariance matrix and vector of mean values after the action of the CP map \mathcal{M} ,

$$\begin{pmatrix} d_{j,k,A} \\ d_{j,k,B} \end{pmatrix} = S \begin{pmatrix} d_{j,k} \\ 0 \end{pmatrix}, \quad \gamma_{j,k,AB} = S(\Gamma_{j,k}^{-1} \oplus I_B)S^T + G. \quad (3.55)$$

We also decompose the inverse matrix $\Gamma_{j,k,AB} = \gamma_{j,k,AB}^{-1}$ similarly as in Eq. (3.33),

$$\Gamma_{j,k,AB} = \begin{bmatrix} \Upsilon_{j,k,A} & \sigma_{j,k} \\ \sigma_{j,k}^T & \Upsilon_{j,k,B} \end{bmatrix}. \quad (3.56)$$

The j -th term in Eq. (3.54) gives rise to two new terms. The $(2j-1)$ -th term is parameterized by

$$\begin{aligned} \Gamma_{2j-1,k} &= \Upsilon_{j,k-1,A} - \sigma_{j,k-1} \Upsilon_{j,k-1,B}^{-1} \sigma_{j,k-1}^T, \\ d_{2j-1,k} &= d_{j,k-1,A} + z_{k+1}, \\ C_{2j-1,k} &= C_{j,k-1}. \end{aligned} \quad (3.57)$$

Similarly, the formulas for the $2j$ -th term read

$$\begin{aligned} \Gamma_{2j,k} &= \Upsilon_{j,k-1,A} - \sigma_{j,k-1} \tilde{\Upsilon}_{j,k-1,B}^{-1} \sigma_{j,k-1}^T, \\ d_{2j,k} &= d_{j,k-1,A} + \Gamma_{2j,k}^{-1} \sigma_{j,k-1} \tilde{\Upsilon}_{j,k-1,B}^{-1} d_{j,k-1,B} + z_{k+1}, \\ C_{2j,k} &= -2C_{j,k-1} \sqrt{\frac{\det(\Gamma_{j,k-1,AB})}{\det(\Gamma_{2j,k}) \det(\tilde{\Upsilon}_{j,k-1,B})}} \\ &\quad \times \exp[-d_{j,k-1,B}^T M d_{j,k-1,B}], \end{aligned} \quad (3.58)$$

where $\tilde{\Upsilon}_{j,k-1,B} = \Upsilon_{j,k-1,B} + I$ and

$$\begin{aligned} M &= \Upsilon_{j,k-1,B} \tilde{\Upsilon}_{j,k-1,B}^{-1} \\ &\quad - \tilde{\Upsilon}_{j,k-1,B}^{-1} \sigma_{j,k-1}^T \Gamma_{2j,k}^{-1} \sigma_{j,k-1} \tilde{\Upsilon}_{j,k-1,B}^{-1}. \end{aligned}$$

Iterating these formulas N times starting from the initial ($k=0$) Gaussian state (3.26) and then applying the final anti-squeezing operation $S^\dagger(s_{\text{out}})$ which acts on the inverse correlation matrices $\Gamma_{j,N}$ and displacements $d_{j,N}$ as in (3.39), one obtains the Wigner function of the conditionally generated state. The probability of state preparation can be calculated simply as the sum $P = \sum_{j=1}^{2^N} C_{j,N}$.

Examples

We shall now consider, as an illustration, the generation of superpositions of $|0\rangle$, $|1\rangle$, and $|2\rangle$. These states, namely,

$$|\psi\rangle = \frac{1}{\sqrt{1 + |c_0|^2 + |c_1|^2}} (c_0|0\rangle + c_1|1\rangle + |2\rangle). \quad (3.59)$$

can be prepared with two photon subtractions. Here, we assume that the coefficient c_2 of the Fock state $|2\rangle$ is non-zero (we arbitrarily take it equal to one). Otherwise, only one (or zero) photon subtraction would be needed to generate the target state. In the case of two photon subtractions interspersed with three displacements, Eq. (3.48) reduces to

$$\begin{aligned} |\psi\rangle_{\text{out}} &\propto (\sinh(s) \cosh(s) + \beta_1 \beta_2) |0\rangle \\ &- (\beta_1 + \beta_2) \sinh(s) |1\rangle + \sqrt{2} \sinh^2(s) |2\rangle. \end{aligned} \quad (3.60)$$

This state matches the target state (3.59) if

$$\beta_{1,2} = \frac{-B \pm \sqrt{B^2 - 4C}}{2}, \quad (3.61)$$

where

$$\begin{aligned} B &= \sqrt{2} \sinh(s) c_1, \\ C &= \sqrt{2} \sinh^2(s) c_0 - \sinh(s) \cosh(s). \end{aligned}$$

Equations (3.49) and (3.50) allow us to calculate the displacements needed to generate this state. Assuming for simplicity that c_0 , c_1 and s are chosen such that β_1 and β_2 are both real, we obtain

$$\begin{aligned} \alpha_3 &= \beta_2, \\ \alpha_2 &= (\beta_1 - \alpha_3)/t, \\ \alpha_1 &= \frac{\tanh(s)(\alpha_3 + \alpha_2/t) - (\alpha_3 + t\alpha_2)}{t^2 - \tanh(s)/t^2}. \end{aligned} \quad (3.62)$$

In order to illustrate our method, let us consider the following four superpositions of the Fock states $|0\rangle$, $|1\rangle$, and $|2\rangle$:

$$|\psi_1\rangle = |2\rangle, \quad (3.63)$$

$$|\psi_2\rangle = \frac{1}{\sqrt{2}}(|1\rangle + |2\rangle), \quad (3.64)$$

$$|\psi_3\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |2\rangle), \quad (3.65)$$

$$|\psi_4\rangle = \frac{1}{\sqrt{3}}(|0\rangle + |1\rangle + |2\rangle), \quad (3.66)$$

We plot the behavior of the fidelity and probability of generation of the target states (3.63) – (3.66) as a function of the initial squeezing s_{in} (Fig. 3.8), beam-splitter transmittance (Fig. 3.9), and photodetector efficiency (Fig. 3.10). As in the preceding section, we observe that the fidelity of the generation for any state gets arbitrarily close to one as T approaches unity, as shown in Fig. 3.9. We also find that the fidelity is very robust against small detector efficiency η , as can be seen in Fig. 3.10. On the other hand, the preparation probability decreases with a growing T and decreasing η , as predicted by the equation $P \propto (1 - T)^2 \eta^2$.

All these features are very similar to those found in the preceding section, where we considered only states generated with one photon subtraction. Let us now stress some new features. First, we note here the existence of a clear optimal input squeezing, giving the maximum fidelity for each of the four studied states, see Fig. 3.8(a). Observing that for a fixed T the optimal squeezing has a higher value [from 2.4 dB for state (3.65) to 4 dB for state (3.64)] than those encountered in the case of one photon subtraction, we can expect an increasing value of the optimal squeezing for an increasing number of

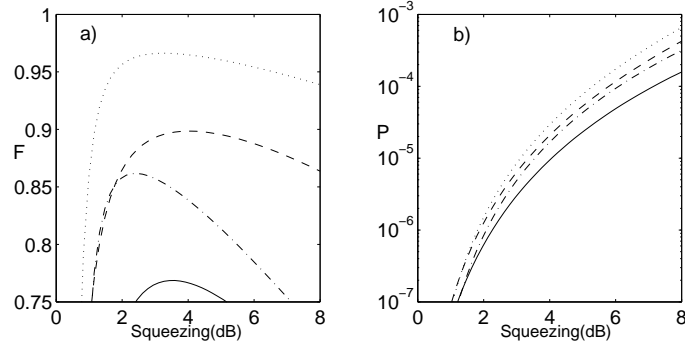


Figure 3.8: (a) Fidelity between the generated state and the target state and (b) probability of successful generation as a function of the squeezing s_{in} for the four target states (3.63) (solid line), (3.64) (dashed line), (3.65) (dot-dashed line), and (3.66) (dotted line), with $T = 0.95$ and $\eta = 0.25$.

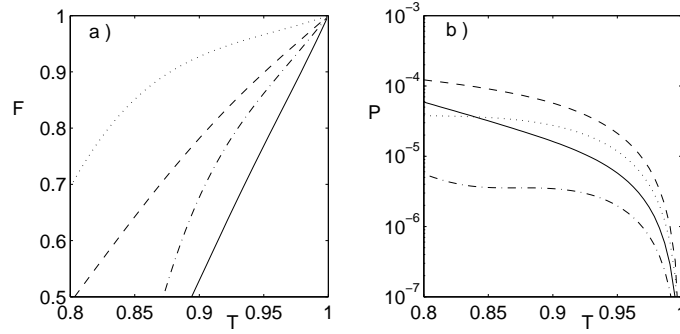


Figure 3.9: (a) Fidelity between the generated state and the target state and (b) probability of successful generation as a function of T for the four target states (3.63)–(3.66). The curves are plotted considering the optimal squeezing s_{in} for each state, namely 3.54 dB for state (3.63) (solid line), 4.02 dB for state (3.64) (dashed line), 2.43 dB for state (3.65) (dot-dashed line), and 3.24 dB for state (3.66) (dotted line). The curves are plotted for $\eta = 0.25$.

Fock states in the target superposition. It can be checked that the value of this optimal input squeezing tends to zero when T tends to 100%, at the expense of a vanishing generation probability.

Another interesting fact is the existence of very different values of the optimum fidelity for different target states for a fixed $T = 0.95$ and $\eta = 0.25$, as shown in Fig. 3.8(a). For example, the two-photon state $|2\rangle$ is much more difficult to generate using our method than the other three states (3.64)–(3.66). For the state $|2\rangle$, a transmittance of $T > 0.99$ is necessary to reach a fidelity of $F > 0.95$, resulting in a very low probability of generation. This would make the experimental generation of $|2\rangle$ (or $S(s_{\text{out}})|2\rangle$ if the final squeezing operation is omitted) with a good fidelity very challenging. In contrast, the balanced superposition state (3.66) can be generated with a high fidelity $F > 0.90$ even with a transmittance $T \approx 0.90$.

Finally, a surprising fact arises when $\beta_1 \neq \beta_2$. Then, the equations (3.62) give two distinct sets of α_i 's generating the same target state, the second set being obtained by making the exchange $\beta_1 \leftrightarrow \beta_2$. Considering the pure-state description and $T \rightarrow 1$, the

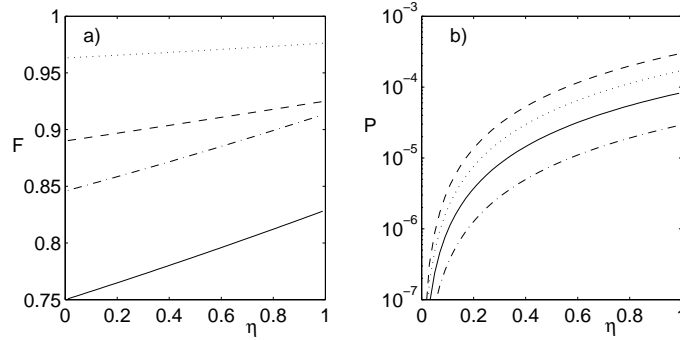


Figure 3.10: (a) Fidelity between the generated state and the target state and (b) probability of successful generation as a function of η for the four target states (3.63)–(3.66). The curves are plotted considering the optimal squeezing s_{in} for each state, namely 3.54 dB for state (3.63) (solid line), 4.02 dB for state (3.64) (dashed line), 2.43 dB for state (3.65) (dot-dashed line), and 3.24 dB for state (3.66) (dotted line). The curves are plotted for $T = 0.95$.

two alternative choices of displacements become strictly equivalent. In contrast, when considering the realistic model with $T < 1$, these two solutions for the same target state do not have exactly the same behavior. As we can see in Fig. 3.11(a), one of the two solutions is indeed more robust to decreasing T . However, the two solutions are rather similar as far as the probability of state generation is concerned, as shown in Fig. 3.11(b).

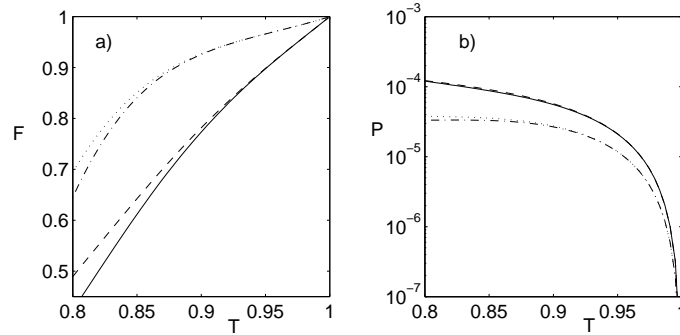


Figure 3.11: (a) Fidelity between the generated state and the target state and (b) probability of successful generation as a function of T for the two target states (3.64) and (3.66). The curves correspond to the two choices of the displacements α_1 and α_2 when considering the optimal squeezing s_{in} for each state, namely 4.02 dB for state (3.64) (dotted line, dot-dashed line), and 3.24 dB for state (3.66) (dashed line, solid line). The curves are plotted for $\eta = 0.25$.

3.4 Efficient State Preparation

We have seen in the preceding sections that the probability of successful state preparation decreases exponentially with the maximum number of photons N in the superposition, $P \propto (1 - T)^N \eta^N$, which would limit the applicability of the scheme to $N \leq 2$ in

practice. In order to overcome this problem and enhance the success rate, we have to use some additional resources besides linear optics and squeezers. The main weakness of the present scheme is that all photon subtractions have to succeed simultaneously for the state to be generated, which results in this exponential scaling. As we will see, this can be avoided provided that a quantum memory is employed. Recently, the first experimental demonstrations of quantum memory for light based on the interaction of light beams with atomic ensembles have been reported [114, 131]. A quantum memory enables to deterministically store the state of a light mode for some time, and to retrieve it later on when required.

Our efficient state preparation scheme works in an iterative manner, with two states with up to $N/2$ photons being generated separately and stored in a quantum memory. The total number of trials required to generate both states then scales as $1/P_A + 1/P_B$, instead of $1/(P_AP_B)$ which would be the case without a memory. The two states are then merged, and a state with up to N photons is produced. This merging is achieved conditionally by combining the two modes on a balanced beam splitter and projecting one of the output modes onto vacuum, see Fig. 3.12(a). This requires an efficient detector, being able to discriminate between the presence and absence of a photon. This is a second extra resource for our efficient state preparation scheme. The scheme is iterative in the sense that each of the states with up to $N/2$ photons is itself obtained by merging two states with up to $N/4$ photons, etc.

The scheme starts from superpositions of vacuum and single photon states $c_0|0\rangle + c_1|1\rangle$, which can be prepared conditionally using the scheme discussed in Section 3.3. These states are repeatedly merged together and after each successful merging the resulting state is stored in a quantum memory, see Fig. 3.12(b). After k successful iterations, an arbitrary state containing up to 2^k photons can be prepared. A similar technique was already proposed in the literature to efficiently generate two-mode N -photon entangled Schrödinger cat-like states [45], and it is inspired by the quantum repeater concept [36, 65] where such a recursive method is exploited to efficiently distribute entanglement through noisy channels over long distances.

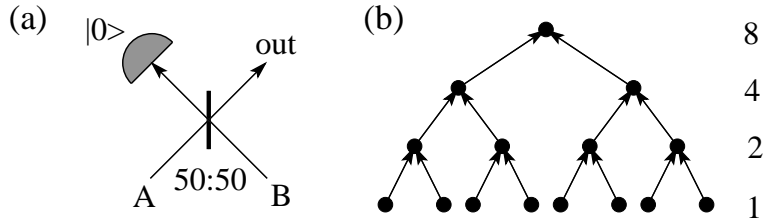


Figure 3.12: (a) Setup for merging two states in modes A and B into a single state. The procedure succeeds conditionally on detecting no photons in the left output mode. (b) Starting from superpositions of $|0\rangle$ and $|1\rangle$ and repeating the merging operation iteratively, it is possible to prepare states with up to 2^n photons after n iterations.

In order to check that the iterative scheme is universal, note that any state $|\psi^{(2N)}\rangle = \sum_{n=0}^{2N} c_n |n\rangle$ can be written as $|\psi^{(2N)}\rangle \propto \prod_{j=1}^{2N} (a^\dagger - \alpha_j) |0\rangle$. Now choose $|\psi_A^{(N)}\rangle \propto \prod_{j=1}^N (\sqrt{2}a_A^\dagger - \alpha_j) |0\rangle$ and $|\psi_B^{(N)}\rangle \propto \prod_{j=N+1}^{2N} (\sqrt{2}a_B^\dagger - \alpha_j) |0\rangle$. If the modes A and B are combined on a balanced beam splitter, then we have

$$\begin{aligned} a_{A,\text{in}} &= \frac{1}{\sqrt{2}}(a_{A,\text{out}} + a_{B,\text{out}}), \\ a_{B,\text{in}} &= \frac{1}{\sqrt{2}}(a_{A,\text{out}} - a_{B,\text{out}}). \end{aligned} \quad (3.67)$$

As a consequence, the state of the output mode A conditioned on projecting B onto vacuum is proportional to $|\psi^{(2N)}\rangle$. This decomposition of $|\psi^{(2N)}\rangle$ into $|\psi_A^{(N)}\rangle$ and $|\psi_B^{(N)}\rangle$ can be repeated until we find the $2N$ basic states $c_0^k|0\rangle + c_1^k|1\rangle$, $k = 1, \dots, 2N$, from which the state $|\psi^{(2N)}\rangle$ can be iteratively prepared.

As a first example, let us consider the preparation of the state $|\psi^{(2)}\rangle = c_0|0\rangle + c_1|1\rangle + c_2|2\rangle$ by merging the states $a_0|0\rangle + a_1|1\rangle$ and $b_0|0\rangle + b_1|1\rangle$, where the coefficients a_j and b_j can be determined by solving a system of equations

$$\frac{b_1}{b_0} + \frac{a_1}{a_0} = \sqrt{2} \frac{c_1}{c_0}, \quad \frac{a_1}{a_0} \frac{b_1}{b_0} = \sqrt{2} \frac{c_2}{c_0}, \quad (3.68)$$

and using the normalization of the states. The probability of successful merging is given by

$$P_M = |a_0|^2 |b_0|^2 + \frac{1}{2} |a_0 b_1 + a_1 b_0|^2 + \frac{1}{2} |a_1|^2 |b_1|^2. \quad (3.69)$$

The probability can be bounded from below, $P_M \geq \frac{1}{3}$, and the minimum is achieved when $a_0 = b_0 = 1/\sqrt{3}$ and $a_1 = -b_1 = \sqrt{2/3}$. It follows that the total number of elementary operations required to generate the state $|\psi^{(2)}\rangle$ is $O_{\text{tot}} \approx 6/P_1$, where P_1 is the probability of preparing the superposition of vacuum and single-photon states. This should be compared with the total number of trials $O_{\text{tot}} \approx 1/P_1^2$ necessary when the scheme described in the preceding section is used instead. As we have seen before, $P_1 \approx 10^{-2}$, hence the present procedure reduces the number of necessary operations by more than an order of magnitude even in this simplest case. The price to pay, of course, is the need for a quantum memory and a highly efficient photo-detector for the merging operation.

In order to show that the required resources scale only sub-exponentially with N , let us consider the preparation of the single-mode states $\frac{1}{\sqrt{2}}(|0\rangle + |N\rangle)$. At the n -th iteration step, states $|0\rangle \pm c_{n-1}|2^{n-1}\rangle$ are merged to produce a state $|0\rangle + c_n|2^n\rangle$ (we omitted the normalization prefactors for simplicity). The coefficients are related as follows,

$$c_n = \frac{c_{n-1}^2}{2^{2^{n-1}-1} 2^{n-1}!} \sqrt{2^{n!}}. \quad (3.70)$$

Starting from $c_{\log_2 N} = 1$ all coefficients c_n , $n < \log_2 N$ can be determined from Eq. (3.70). The probability of successful merging is given by

$$P_{(n-1) \rightarrow n} = \frac{1 + |c_n|^2}{(1 + |c_{n-1}|^2)^2}, \quad (3.71)$$

and the total number of operations to prepare the state $|0\rangle + c_n|2^n\rangle$ can be estimated as $O_n = 2O_{n-1}/P_{n-1 \rightarrow n}$. For large $K = 2^n$, we can use the Stirling approximation $K! \approx \sqrt{2\pi K} K^K e^{-K}$ and we get $c_n \approx c_{n-1}^2/(\pi 2^{n-1})^{1/4}$. Within this approximation, we can bound the probability (3.71) as follows,

$$P_{n-1 \rightarrow n} \approx \frac{1}{\sqrt{\pi 2^{n-1}}} \frac{\sqrt{\pi 2^{n-1}} + |c_{n-1}|^4}{(1 + |c_{n-1}|^2)^2} \geq \frac{1}{2\sqrt{\pi 2^{n-1}}}. \quad (3.72)$$

The recurrence formula for the total number of operations becomes $O_n = 2\sqrt{\pi 2^{n+1}} O_{n-1}$ which can be solved to yield

$$O_n = \frac{1}{P_1} (2\sqrt{2\pi})^n 2^{n(n+1)/4}, \quad (3.73)$$

where P_1 is the probability of preparation of $|0\rangle + c_0|1\rangle$. An approximate bound on the total number of operations $O_{\text{cat,tot}}(N)$ required to generate the state $\frac{1}{\sqrt{2}}(|0\rangle + |N\rangle)$

can be obtained from Eq. (3.73) by setting $n = \log_2 N$ and we get

$$O_{\text{cat,tot}}(N) \leq \frac{1}{P_1} N^{\frac{7}{4} + \frac{1}{2} \log_2 \pi} N^{\frac{1}{4} \log_2 N}, \quad (3.74)$$

which is clearly a sub-exponential scaling with N . In Fig. 11, we plot the total number of operations as a function of N determined by numerical calculations. The log-log plot reveals that the dependence of $O_{\text{cat,tot}}(N)$ on N is essentially polynomial.

As an example, consider the case $N = 8$. Assuming $P_1 = 10^{-2}$ we get $O_{\text{cat,tot}}(8) = 37000$, while if using the scheme discussed in Section 3.3 then eight photon subtractions would have to be performed simultaneously and about $(1/P_1)^8 = 10^{16}$ trials would be required. The scheme employing a quantum memory is thus eleven orders of magnitude more efficient.

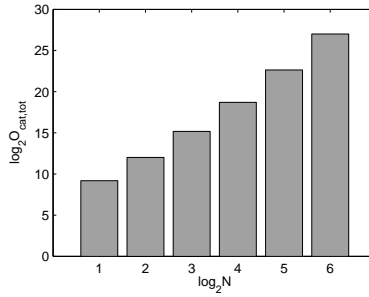


Figure 3.13: Total number of operations required to prepare a single-mode Schrödinger cat-like state $(|0\rangle + |N\rangle)/\sqrt{2}$.

3.5 Conclusions

In summary, we have shown that an arbitrary single-mode state of light can be engineered starting from a squeezed vacuum state and applying a sequence of displacements and single-photon subtractions, followed by a final squeezing operation. Since our proposal does not require single-photon sources and can operate with low-efficiency photodetectors, we anticipate that its experimental implementation will be much easier than for the previous proposals, in particular the one based on repeated photon additions [55].

We have shown that an arbitrary superposition of $|0\rangle$, $|1\rangle$ and $|2\rangle$ states can be successfully produced with high fidelity using a reasonably low squeezing ($\simeq 3\text{dB}$) if the transmittance T of the beam splitter used for photon subtraction is sufficiently close to unity (e.g. $T \simeq 95\%$). This holds even when inefficient photodetectors with single-photon sensitivity but no single-photon resolution are employed, such as the standard avalanche photodiodes, as it only affects the probability of successful generation of the state without compromising the fidelity. However, low η and high T drastically reduce the preparation probability, so that a compromise has to be made when determining T . The final anti-squeezing operation required to obtain a finite superposition of Fock states is technically perhaps the most demanding part of the scheme, but is nevertheless achievable with the current technology.

The recent demonstrations of single-photon subtraction from a single-mode squeezed vacuum [142, 197] provides a strong evidence of the practical feasibility of our scheme. We may reasonably assert, based on our proposal, that the preparation of squeezed superpositions of $|0\rangle$, $|1\rangle$, and $|2\rangle$ states should be experimentally achievable with the present technology.

Chapter 4

Loophole-free Bell Test

4.1 Introduction

Since the inception of quantum mechanics, several physicists have considered its counterintuitive aspect as an evidence of the incompleteness of the theory. There have been repeated suggestions that its probabilistic features may possibly be described by an underlying deterministic substructure. The first attempt in this direction originates from the famous paper by Einstein, Podolsky, and Rosen (EPR) [69] in 1935. There, it was advocated that if “local realism” (causality + deterministic substructure, as described below) is taken for granted, then quantum theory is an incomplete description of the physical world.

The EPR argument gained a renewed attention in 1964, when John Bell derived his famous inequalities, which must be satisfied within the framework of any local realistic theory [15]. Bell showed that any such deterministic substructure model (also called “hidden-variables model”), if local, yields predictions that significantly differ from those of quantum mechanics. The merit of Bell inequalities lies in the possibility to test them experimentally, allowing physicists to test whether either quantum mechanics or local realism is the correct description of Nature.

Bell Inequalities

In this chapter, we will use the Clauser-Horne-Shimony-Holt inequality (called Bell-CHSH inequality in the following), originally devised for a two-qubit system [48]. Let us consider the following thought experiment, which we will analyze from the point of view of local realism. The experiment involves three distant parties, Sophie, Alice, and Bob. Sophie (the source) prepares a bipartite state and distribute it to Alice and Bob (the two usual partners), see Fig. 4.1.

Then, Alice and Bob randomly and independently decide between one of two possible quantum measurements A_1 or A_2 (B_1 or B_2), which should have only two possible outcomes $+1$ or -1 . The timing of the experiment should be arranged in such a way that Alice and Bob do their measurements in a causally disconnected manner. Thereby, Alice’s measurement cannot influence Bob’s, and vice-versa. Local realism implies two assumptions:

1. Realism: the physical properties A_1, A_2, B_1, B_2 have definite values a_1, a_2, b_1, b_2 , which exist independently of their observation. This implies the existence of a probability distribution $P(a_1, a_2, b_1, b_2)$, dependent on how Sophie generates the bipartite state.

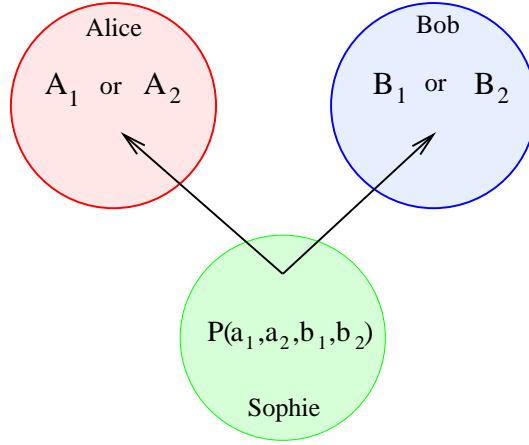


Figure 4.1: Sophie prepares a bipartite state and distributes it to Alice and Bob, who perform each a measurement. Alice measures either A_1 or A_2 , while Bob measures B_1 or B_2 . In a local realistic theory, there must exist an underlying probability distribution $p(a_1, a_2, b_1, b_2)$, generated by Sophie.

2. Locality: Alice's measurement choice and outcome do not influence the result of Bob's measurement, and vice-versa. The measurement events are separated by a spacelike interval.

If we consider local realism as the correct description of the physical world, then we obtain the Bell-CHSH inequality

$$|S| = |\langle a_1 b_1 \rangle + \langle a_1 b_2 \rangle + \langle a_2 b_1 \rangle - \langle a_2 b_2 \rangle| \leq 2, \quad (4.1)$$

where $\langle a_j b_k \rangle$ denotes the average over the subset of experimental data where Alice measured a_j and, simultaneously, Bob measured b_k . Indeed, if there is an underlying probability distribution $p(a_1, a_2, b_1, b_2)$, then each realization of it contributes by $a_1(b_1 + b_2) + a_2(b_1 - b_2) = \pm 2$ to the average, implying Eq. (4.1).

Now, if we consider that Sophie generates and distributes an entangled pair of qubits, quantum mechanics predicts $S = 2\sqrt{2}$, which is in contradiction with local realism. Thus, an experimental test of Bell-CHSH inequalities where a violation of $S \leq 2$ is observed disproves any classical (local realistic) description of Nature.

Experimental Bell Test and Related Loopholes

From the beginning of the 80's, many experimental Bell tests [8, 9, 10, 80, 122, 188, 195] have been performed, observing the violation of Bell inequalities predicted by quantum mechanics. All these schemes used optical setups because, at that time, it was the only known way of generating and distributing entangled particles (photons) at a distance in order to make Alice's and Bob's measurements causally disconnected. The technology of generation of entangled states of photons is very well mastered today [122] and the prepared entangled states can be distributed over long distances via low-loss optical fibers [195]. However, the currently available single-photon detectors continue to suffer from a too low efficiency η_{PD} , which can be exploited by a local realistic model to yield a violation. Thus, to reject local realism, it is necessary to make the extra assumption that the registered pairs form a fair sample of the emitted pairs. So, from a logical point of view, these experiments do not succeed in ruling out a local realistic model; this is the so-called *detector-efficiency loophole* [121, 146, 163]. This loophole has been

closed in a recent experiment with trapped ions [161], thanks to the high efficiency of the measurement of the ion states. However, the ions were held in a single trap, only several micrometers apart, so that the measurement events were not spacelike separated, opening in turn the so-called *locality loophole* [16, 164].

So far, no experimental test has succeeded to close both loopholes at the same time, that is, the measured correlations may be explained in terms of local realistic theories exploiting the low detector efficiency or the timelike interval between the two detection events. It was suggested that two distant trapped ions can be entangled via entanglement swapping by first preparing an entangled state of an ion and a photon on each side and then projecting the two photons on a maximally entangled singlet state [37, 64, 75, 182]. Very recently, the first step toward this goal, namely the entanglement between a trapped ion and a photon emitted by the ion, has been observed experimentally [24]. However, the entanglement swapping would require interference of two photons emitted by two different ions, which is experimentally very challenging. An interesting alternative to the atom-based approaches [81, 82, 182] consists of all-optical schemes based on continuous variables of light. Indeed, the balanced homodyne detection used in these schemes can exhibit a high detection efficiency [152], sufficient to close the detection loophole.

4.2 Bell Test with Continuous Variables of Light

Quantum continuous variables of light have been successfully used to realize some of the standard informational tasks traditionally based on qubits. Unfortunately, the entangled two-mode squeezed state that can easily be generated experimentally [32, 140, 167] cannot be directly employed to test Bell inequalities with homodyning. Indeed, as noted by Bell himself, this state is described by a positive-definite Gaussian Wigner function, which thus provides a local realistic model that can explain all correlations between quadrature measurements (carried out by balanced homodyne detectors). Thus, similarly as in the case of the purification of continuous variable entanglement [37, 71, 73, 79, 89], one has to go beyond the class of Gaussian states or Gaussian operations.

In particular, it is possible to obtain a Bell violation with a Gaussian two-mode squeezed vacuum state by performing a non-Gaussian measurement, for example a photon-counting measurement [13]. As shown in Fig. 4.2, Sophie prepares an entangled state and distributes it to Alice and Bob. The two possible measurements on Alice's and Bob's sides consist in randomly choosing between applying the displacement $D(\alpha)$ or no displacement, followed by a measurement of the parity of the number of photons n impinging on the single-photon detector. The resulting parity $a_i = (-1)^n$ gives the binary result used in the Bell-CHSH inequality. It can be shown (see [13]) that

$$S = |W(0, 0) + W(\alpha, 0) + W(0, \alpha) - W(\alpha, \alpha)| \quad (4.2)$$

where $W(x, p)$ is the Wigner function of the entangled state, violates the Bell-CHSH inequality $S \leq 2$ by about 10% for an appropriate choice of α . Recent proposals using more abstract measurements described in Refs. [46, 76, 113] gave similar results. Note, however, that these measurements are either experimentally infeasible or suffer from a very low detection efficiency, thereby re-opening the detection loophole.

Considering the current state of the art in quantum optics technologies, the scheme based on high-efficiency homodyne detection seems to be the most promising way of closing the detection loophole. However, since homodyning is a Gaussian measurement, it is then necessary to generate highly non-classical non-Gaussian entangled states, whose Wigner function is not positive definite. In addition, one has to develop a

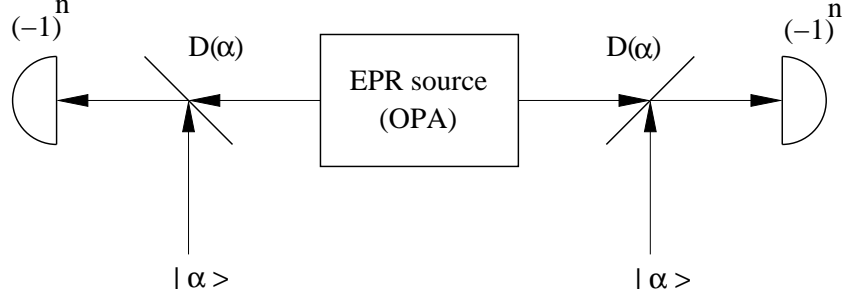


Figure 4.2: Bell test using the parity of the number of photons impinging on each photodetector. Sophie prepares an entangled state (EPR) and distributes it to Alice and Bob. Each of them either applies a displacement $D(\alpha)$ or not, and uses the parity of the number of photons measured using a photodetector with single-photon resolution [13].

method for converting the continuous result obtained by homodyne measurement into a binary result (the so-called “binning” method).

Several recent theoretical works have demonstrated that a violation of Bell inequalities can be observed using balanced homodyning provided that specific entangled light states such as pair-coherent states, squeezed Schrödinger cat-like states, or specifically tailored finite superpositions of Fock states, are available [90, 91, 134, 196]. More specifically, the violation of the Bell-CHSH inequality was derived in Ref. [134] for a state of the form

$$|\psi_{\text{in}}\rangle_{AB} = \sum_{n=0}^{\infty} c_n |n, n\rangle_{AB}, \quad (4.3)$$

with $|n\rangle$ denoting Fock states, and a binning based on the sign of the measured quadrature. Optimizing over the quadrature angles and probability amplitudes c_n (see Fig. 4.3), one obtains a maximal Bell-CHSH inequality violation of $S = 2.076$. Interestingly, it was shown in Ref. [196] that the highest possible violation of $S = 2\sqrt{2}$ can be obtained with the bipartite state

$$|\psi_{\text{in}}\rangle_{AB} = |f, f\rangle + e^{i\theta} |g, g\rangle, \quad (4.4)$$

where $f(q)$ and $g(q)$ are the wave functions of some specific states, and a more complicated binning based on the roots of $f(q)$ and $g(q)$ is used. Unfortunately, no feasible experimental scheme is known today that could generate the states required in Refs. [90, 91, 134, 196].

In a recent experiment, an entangled state obtained by splitting a single photon on a balanced beam splitter was used to make a Bell test where an homodyne detection was carried out on each output [11]. It was claimed that the observed data violate Bell inequality; however, the violation was obtained by post-selecting only the data when the absolute value of the detected quadrature was above some threshold. This rejection of data introduces a loophole very similar to the detection efficiency loophole, and this experiment therefore does not refute local realism.

Recently, we showed [85] (and independently by Nha and Carmichael [135]), that a very simple non-Gaussian state obtained by subtracting a single photon from each mode of a two-mode squeezed vacuum state can exhibit a Bell violation with homodyning. Note that this non-Gaussian state is close to the optimal state obtained in Ref. [134], as is visible in Fig. 4.3, and gives a violation of $S = 2.046$. An essential feature of this proposal is that the photon subtraction can be successfully performed with low-efficiency single-photon detectors [49, 138, 139], which renders the setup experimentally

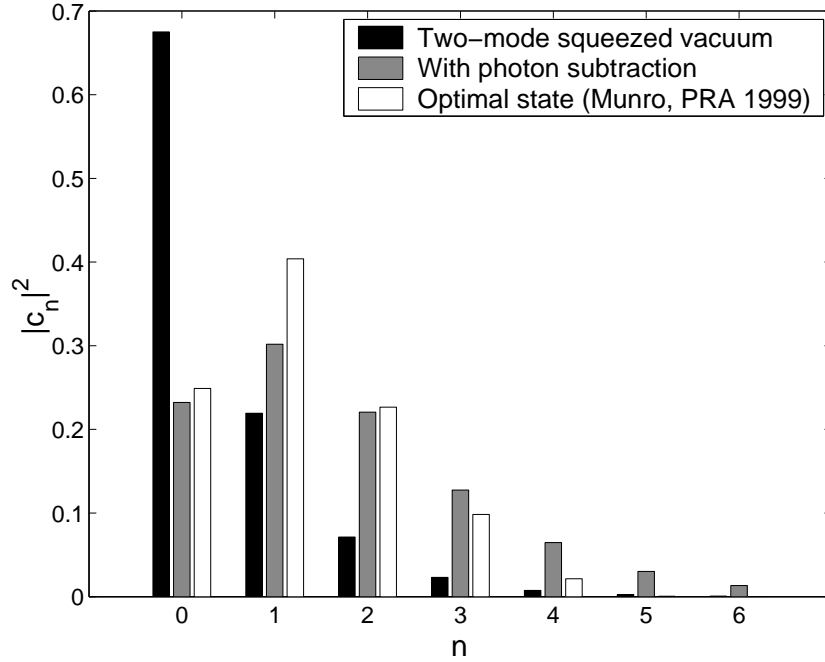


Figure 4.3: Probabilities $|c_n|^2$ in the Fock basis of the two-mode squeezed vacuum state with $\lambda = 0.57$ (black), the non-Gaussian state obtained from the previous state by subtracting one photon from each mode (grey), and the optimal state of Ref. [134] (white).

feasible. In fact, the basic building block of the scheme, namely the de-Gaussification of a single-mode squeezed vacuum via single-photon subtraction, has recently been demonstrated experimentally [197]. In the following, we provide a thorough analysis of the scheme proposed in Refs. [85, 135]. We present the details of the calculation of the Bell factor for a realistic setup that takes into account mixed input states, losses, added noise and imperfect detectors. Moreover, we shall also discuss several alternative schemes that involve the subtraction of one, two, three, or four photons.

4.3 Feasible Bell Test with Homodyne Detection

Proposed Optical Setup

The conceptual scheme of the proposed experimental setup is depicted in Fig. 4.4. A source generates a two-mode squeezed vacuum state in modes A and B. This can be accomplished, e.g., by means of non-degenerate parametric amplification in a $\chi^{(2)}$ nonlinear medium or by generating two single-mode squeezed vacuum states and combining them on a balanced beam splitter. Subsequently, the state is de-gaussified by conditionally subtracting a single photon from each beam. A tiny part of each beam is reflected from a beam splitter BS_A (BS_B) with a high transmittance T . The reflected portions of the beams impinge on single-photon detectors such as avalanche photodiodes. A successful photon number subtraction is heralded by a click of each photodetector PD_A and PD_B [138]. In practice, the photodetectors exhibit a single-photon sensitivity but not a single photon resolution, that is, they can distinguish the absence and presence of photons but cannot measure the number of photons in the

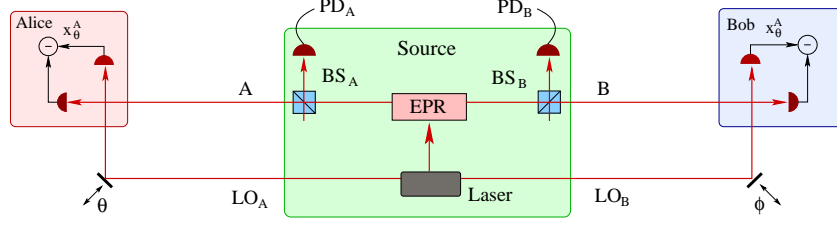


Figure 4.4: Conceptual scheme of the proposed experimental setup for observing a violation of Bell inequalities with balanced homodyne detection. The source emits a two-mode squeezed vacuum state in modes A and B. A small part of the beams is subtracted on two unbalanced beam splitters BS_A and BS_B and sent on single-photon detectors PD_A and PD_B . The two remaining beams A and B, which are conditionally prepared in a non-Gaussian entangled state, are sent to Alice and Bob, respectively, who perform each a balanced homodyne detection using their local oscillator LO_A and LO_B .

mode. Nevertheless, this is not a problem here because in the limit of high T , the most probable event leading to the click of a photodetector is precisely that a single photon has been reflected from the squeezed beam on the beam splitter. The probability of an event where two or more photons are subtracted from a single mode is smaller by a factor of $\approx 1 - T$ and becomes totally negligible in the limit of $T \rightarrow 1$. Another important feature of the scheme is that the detector efficiency η_{PD} can be quite low because small η_{PD} only reduces the success rate of the conditional single-photon subtraction but it does not significantly decrease the fidelity of this operation. These issues will be discussed in detail in Section 4.4.

After generation of the non-Gaussian state, the two beams A and B together with the appropriate local oscillators LO_A and LO_B are sent to Alice and Bob, who then randomly and independently measure one of two quadratures $\hat{x}_{\theta_j}^A, \hat{x}_{\phi_k}^B$ characterized by the relative phases θ_1, θ_2 and ϕ_1, ϕ_2 between the measured beam and the corresponding local oscillator. The rotated quadratures $\hat{x}_{\theta}^A = \cos \theta \hat{x}^A + \sin \theta \hat{p}^A$ and $\hat{x}_{\phi}^B = \cos \phi \hat{x}^B + \sin \phi \hat{p}^B$ are defined in terms of the four quadrature components of modes A and B that satisfy the canonical commutation relations $[\hat{x}^j, \hat{p}^k] = i\delta_{jk}, j, k \in \{A, B\}$.

Proposed Binning

In the proposed experiment, Alice and Bob measure quadratures which have continuous spectrum. We discretize the quadratures by postulating that the outcome is $+1$ when $x \geq 0$ and -1 otherwise. The two different measurements on each side correspond to the choices of two relative phases θ_1, θ_2 and ϕ_1, ϕ_2 . Quantum mechanically, the correlation $E(\theta_j, \phi_k) \equiv \langle a_j b_k \rangle$ can be expressed as

$$E(\theta_i, \phi_k) = \int_{-\infty}^{\infty} \text{sign}(x_{\theta_i}^A x_{\phi_k}^B) P(x_{\theta_i}^A, x_{\phi_k}^B) dx_{\theta_i}^A dx_{\phi_k}^B, \quad (4.5)$$

where $P(x_{\theta_i}^A, x_{\phi_k}^B) \equiv \langle \hat{x}_{\theta_i}^A, \hat{x}_{\phi_k}^B | \rho_{c,AB} | \hat{x}_{\theta_i}^A, \hat{x}_{\phi_k}^B \rangle$ is the joint probability distribution of the two commuting quadratures $\hat{x}_{\theta_i}^A$ and $\hat{x}_{\phi_k}^B$, and $\rho_{c,AB}$ denotes the (normalized) conditionally generated non-Gaussian state of modes A and B. In practice, the correlations would be determined from the subset of the experimental data corresponding to the successful conditional de-Gaussification, i.e., Alice and Bob would discard all results obtained in measurement runs where either PD_A or PD_B did not click. We emphasize again that this does not open any loophole in the Bell test.

Avoiding the Locality Loophole

To avoid the locality loophole, the whole experiment has to be carried out in the pulsed regime and a proper timing is necessary. In particular, the measurement events on Alice's and Bob's sides (including the choice of phases) have to be space-like separated. A specific feature of the proposed setup is that the non-Gaussian entangled state needed in the Bell test is generated conditionally when both "event-ready" detectors [16] PD_A and PD_B click. However, we would like to stress that this does not represent any loophole if proper timing is satisfied. Namely, in each experimental run, the detection of the clicks (or no-clicks) of photodetectors PD_A and PD_B at the source should be space-like separated from Alice's and Bob's measurements. This guarantees that the choice of the measurement basis on Alice's and Bob's sides cannot in any way influence the conditioning "event-ready" measurement [16, 85, 182]).

As we shall see, exploiting the fact that PD_A and PD_B can be viewed here as "event-ready" detectors [16], one can prove that all local-realistic models for Alice and Bob measurements must satisfy the Bell-CHSH inequality $|S| \leq 2$. In the formalism of "event-ready" detectors introduced by John Bell [16], one should know, by some initiating event, when a measurable system has been produced. The main idea is to pre-select – rather than post-select – the relevant events. For that purpose, one considers three partners, Alice and Bob who perform the measurements, and Sophie who controls the source, see Fig. 4.4. The entire data analysis must be performed on a pulsed basis, with Sophie sending time-tagged light pulses (local oscillator and squeezed light) to Alice and Bob. In each experimental run, Sophie records whether her photodetectors clicked, while Alice and Bob carry out spacelike separated measurements of one of two randomly chosen quadratures. After registering a large number of events, the three partners discard all events not corresponding to an "event-ready" double-click registered by Sophie. The correlation coefficients $\langle a_j b_k \rangle$ are then evaluated from all remaining events, and plugged into the S parameter (4.1). In a local realistic approach, the light pulses in each time slot supposedly carry some random unknown parameter μ , which ultimately determines the sign of \hat{x}_θ^A and \hat{x}_ϕ^B . Imposing by proper timing that the clicks of Sophie's conditioning detectors cannot be influenced by the measurements on Alice's and Bob's sides implies that the probability distribution $p(\mu)$ is independent of the measurement phases $\theta_{1,2}$ and $\phi_{1,2}$. The measured sign s_A on Alice's side (resp. s_B on Bob's side) therefore only depends on μ and θ (resp. ϕ on Bob's side), so that $\langle a_j b_k \rangle = \int d\mu p(\mu) s_A(\theta_j, \mu) s_B(\phi_k, \mu)$, from which the derivation of the Bell-CHSH inequality is very standard [7]. Consequently, a truly "loophole-free" Bell test can be performed provided that Sophie's "event-ready" detectors effectively pre-select the measuring events.

A scheme for observing a Bell-CHSH inequality violation with balanced homodyne very similar to the setup depicted in Fig. 4.4 was proposed by Nha and Carmichael [135]. They also consider de-Gaussification by means of photon subtraction with inefficient detectors exhibiting single-photon sensitivity but no single-photon resolution. The difference between the setup shown in Fig. 4.4 and the scheme of Nha and Carmichael is that in the latter case the single-photon detectors are located on Alice's and Bob's sides while in our case the detectors are spatially separated from the two observers. The position of the photodetectors is irrelevant as far as the state preparation is concerned and both schemes conditionally produce the same photon subtracted two-mode squeezed vacuum. However, the position of these detectors plays a crucial role in the Bell test. If the single-photon detectors are placed together with the balanced homodyne detectors on Alice's and Bob's sides, then the choice of the measurement basis may influence (within the local-hidden-variable models) whether the single-photon detector will click or not. This must be avoided, which is achieved by spatially separating the state preparation and homodyne detection and by proper timing as in our setup.

Ideal Photodetectors

We shall first present a simplified description of the setup, assuming ideal photodetectors ($\eta_{PD} = 1$) with single-photon resolution and conditioning on detecting exactly a single photon at each detector [49, 139]. This idealized treatment is valuable since it provides an upper bound on the practically achievable Bell factor S . Moreover, as noted above, in the limit of high transmittance of BS_A and BS_B , $T \rightarrow 1$, the realistic (inefficient) detector with single-photon sensitivity is in our case practically equivalent to these idealized detectors.

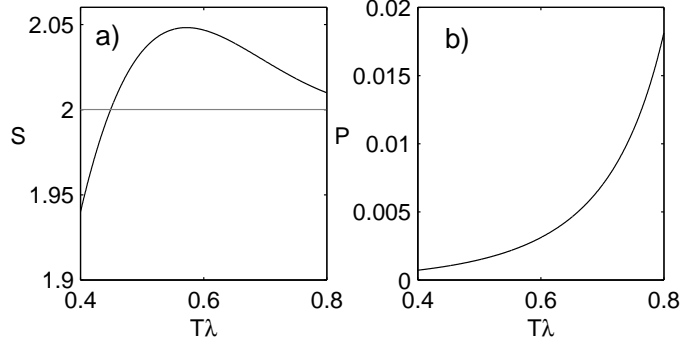


Figure 4.5: (a) Bell factor S is plotted as a function of the effective squeezing parameter $T\lambda$ for $\theta_1 = 0$, $\theta_2 = \pi/2$, $\phi_1 = -\pi/4$ and $\phi_2 = \pi/4$. (b) Probability P of successful conditional generation of the state $|\psi_{\text{out}}\rangle$ as a function of the effective squeezing parameter $T\lambda$, assuming $T = 0.95$.

The two-mode squeezed vacuum state can be expressed in the Fock state basis as follows,

$$|\psi_{\text{in}}(\lambda)\rangle_{AB} = \sqrt{1 - \lambda^2} \sum_{n=0}^{\infty} \lambda^n |n, n\rangle_{AB}, \quad (4.6)$$

where $\lambda = \tanh(s)$ and s is the squeezing constant. In the case of ideal photodetectors, the single-photon subtraction results in the state

$$|\psi_{\text{out}}\rangle_{AB} \propto \hat{a}_A \hat{a}_B |\psi_{\text{in}}(T\lambda)\rangle_{AB}, \quad (4.7)$$

where $\hat{a}_{A,B}$ are annihilation operators and the parameter λ is replaced by $T\lambda$ in order to take into account the transmittance of BS_A and BS_B . A detailed calculation (see Appendix F) yields

$$|\psi_{\text{out}}\rangle_{AB} = \sqrt{\frac{(1 - T^2\lambda^2)^3}{1 + T^2\lambda^2}} \sum_{n=0}^{\infty} (n+1)(T\lambda)^n |n, n\rangle_{AB}, \quad (4.8)$$

and the probability of the conditional preparation of state (4.8) can be expressed as

$$\mathcal{P} = (1 - T)^2 \lambda^2 (1 - \lambda^2) \frac{1 + T^2 \lambda^2}{(1 - T^2 \lambda^2)^3}. \quad (4.9)$$

For pure states exhibiting perfect photon-number correlations, the correlation coefficient (4.5) depends only on the sum of the angles, $E(\theta_j, \phi_k) = \mathcal{E}(\theta_j + \phi_k)$. With the

help of the general formula derived by Munro [134] we obtain for the state (4.8)

$$\begin{aligned} \mathcal{E}(\varphi) = & \frac{(1 - T^2\lambda^2)^3}{1 + T^2\lambda^2} \sum_{n>m} \frac{8\pi(2T\lambda)^{n+m}}{n!m!(n-m)^2} (n+1)(m+1) \\ & \times [\mathcal{F}(n, m) - \mathcal{F}(m, n)]^2 \cos[(n-m)\varphi], \end{aligned} \quad (4.10)$$

where $\mathcal{F}^{-1}(n, m) = \Gamma((1-n)/2)\Gamma(-m/2)$ and $\Gamma(x)$ stands for the Euler gamma function.

We have numerically optimized the angles $\theta_{1,2}$ and $\phi_{1,2}$ to maximize the Bell factor S . It turns out that for any λ , it is optimal to choose $\theta_1 = 0$, $\theta_2 = \pi/2$, $\phi_1 = -\pi/4$ and $\phi_2 = \pi/4$. The Bell factor S for this optimal choice of angles is plotted as a function of the effective parameter $T\lambda$ in Fig. 4.5(a), and the corresponding probability of success of the conditional preparation of the state $|\psi_{\text{out}}\rangle$ is plotted in Fig. 2(b). We can see that S is higher than 2 so the Bell inequality is violated when $T\lambda > 0.45$. The maximal violation is achieved for $T\lambda \approx 0.57$, giving $S \approx 2.048$. This figure is quite close to the maximum Bell factor $S = 2.076$ that could be reached with homodyne detection, sign binning, and arbitrary states exhibiting perfect photon-number correlations $|\psi\rangle = \sum_n c_n |n, n\rangle$ [134].

4.4 Realistic Model

In this section we will consider a realistic scheme with inefficient ($\eta_{\text{PD}} < 1$) photodetectors exhibiting single photon sensitivity but no single-photon resolution, and realistic homodyning with efficiency $\eta_{\text{BHD}} < 1$. The mathematical description of this realistic model of the proposed experiment becomes strikingly simple if we work in the phase-space representation and use the Wigner function formalism. Even though the state used to test Bell inequalities is non-Gaussian, it can be expressed as a linear combination of four Gaussian states, so all the powerful Gaussian tools can still be used.

Two Photon Subtractions

We shall now present a detailed calculation of the Bell factor for our proposed setup, taking into account realistic photodetectors ($\eta_{\text{PD}} < 1$) with single-photon sensitivity (but not resolution), imperfect homodyning, and added electronics noise. The calculation is an extension to two modes states of the phot subtraction presented in the previous chapter.

Preparation of a Non-Gaussian State

As shown in Fig. 4.6, the modes A and B are initially prepared in a two-mode squeezed vacuum state, and the auxiliary modes C and D are in vacuum state. The Wigner function of the four-mode state ABCD is a Gaussian centered at the origin,

$$W_{\text{in},ABCD} = \frac{\sqrt{\det \Gamma_{\text{in}}}}{\pi^4} \exp[-r^T \Gamma_{\text{in}} r], \quad (4.11)$$

where $r = [x^A, p^A, \dots, x^D, p^D]$, and $\Gamma_{\text{in}} = \gamma_{\text{in}}^{-1}$. The initial state is fully characterized by the covariance matrix

$$\gamma_{\text{in}} = \gamma_{\text{TMS},AB} \oplus I_{CD}, \quad (4.12)$$

where γ_{TMS} is the covariance matrix of a two-mode squeezed vacuum and \oplus denotes the direct sum of matrices.

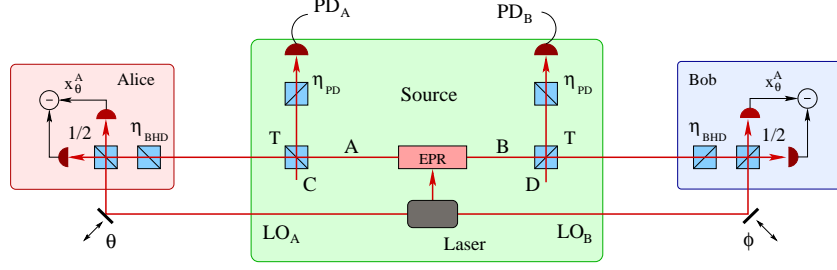


Figure 4.6: Scheme of the proposed experimental setup for observing a violation of Bell inequalities considering realistic photodetectors ($\eta_{\text{APD}} < 1$) with single-photon sensitivity, imperfect homodyning ($\eta_{\text{BHD}} < 1$), and unbalanced beam splitters of transmittance $T < 1$.

The imperfect single-photon detectors (balanced homodyne detectors) with detector efficiency η_{PD} (η_{BHD}) are modeled as a sequence of a lossy channel with transmittance η_{PD} (η_{BHD}) followed by an ideal photodetector (homodyne detector). In our setup, the modes AC (BD) interfere on the unbalanced beam splitters BS_A (BS_B) and pass through the four “virtual” lossy channels before impinging on ideal detectors. The covariance matrix of the mixed Gaussian state $\rho_{\text{out},ABCD}$ just in front of the (ideal) detectors is related to γ_{in} via a Gaussian CP map,

$$\gamma_{\text{out}} = S_{\eta} S_{\text{mix}} \gamma_{\text{in}} S_{\text{mix}}^T S_{\eta}^T + G, \quad (4.13)$$

where

$$S_{\eta} = \sqrt{\eta_{\text{BHD}}} I_{AB} \oplus \sqrt{\eta_{\text{PD}}} I_{CD}, \quad (4.14)$$

$$G = (1 - \eta_{\text{BHD}}) I_{AB} \oplus (1 - \eta_{\text{PD}}) I_{CD}, \quad (4.15)$$

and the symplectic matrix

$$S_{\text{mix}} = S_{BS,AC} \oplus S_{BS,BD} \quad (4.16)$$

describes the mixing of modes A with C and B with D on the unbalanced beam splitters BS_A and BS_B , respectively.

The state $\rho_{c,AB}$ is prepared by conditioning on observing clicks at both photodetectors PD_A and PD_B . These detectors respond with two different outcomes, either a click, or no click. Mathematically, an ideal detector with a single photon sensitivity is described by a two-component positive operator valued measure (POVM) consisting of the projectors onto the vacuum state and on the rest of the Hilbert space, $\Pi_0 = |0\rangle\langle 0|$, $\Pi_1 = I - |0\rangle\langle 0|$. The resulting conditionally prepared state $\rho_{c,AB}$ can be calculated from the density matrix $\rho_{\text{out},ABCD}$ as follows,

$$\rho_{c,AB} = \text{Tr}_{CD}[\rho_{\text{out},ABCD}(I_{AB} \otimes \Pi_{1,C} \otimes \Pi_{1,D})]. \quad (4.17)$$

It is instructive to rewrite the partial trace in Eq. (4.17) in terms of Wigner functions, taking into account that

$$\text{Tr}[XY] = (2\pi)^N \int_{-\infty}^{\infty} W_X(r) W_Y(r) d^{2N}r, \quad (4.18)$$

where $W_X(r)$ and $W_Y(r)$ denote the Wigner representations of the operators X and Y , respectively, and N is the number of modes we trace over. The POVM element Π_1 is a

difference of two operators whose Wigner representations are both Gaussian functions, $W_I = 1/(2\pi)$, $W_0 = \pi^{-1}e^{-x^2-p^2}$. After a bit lengthy but otherwise straightforward calculations, as shown in the previous chapter, we find that the Wigner function $W_{c,AB}$ of (normalized) conditionally prepared state (4.17) can be expressed as a linear combination of four Gaussian functions,

$$W_{c,AB}(r) = \frac{\sqrt{\det \Gamma_{\text{out}}}}{\pi^2 P_G} \sum_{j=1}^4 \frac{q_j}{\sqrt{\det \Gamma_{j,CD}}} e^{-r^T \Gamma_{j,AB} r}, \quad (4.19)$$

where $q_1 = 1$, $q_2 = q_3 = -2$ and $q_4 = 4$. The corresponding probability of success is given by

$$P_G = \sqrt{\det \Gamma_{\text{out}}} \sum_{j=1}^4 \frac{q_j}{\sqrt{\det(\Gamma_{j,AB} \Gamma_{j,CD})}}. \quad (4.20)$$

To define the various matrices appearing in Eqs. (4.19) and (4.20), we first introduce a matrix $\Gamma = \gamma_{\text{out}}^{-1}$ and we divide Γ into four smaller submatrices with respect to the bipartite AB vs CD splitting,

$$\Gamma = \begin{bmatrix} \Gamma_{AB} & \sigma \\ \sigma^T & \Gamma_{CD} \end{bmatrix}. \quad (4.21)$$

It holds that

$$\Gamma_{j,AB} = \Gamma_{AB} - \sigma \Gamma_{j,CD}^{-1} \sigma^T, \quad (4.22)$$

and the four matrices $\Gamma_{j,CD}$ read

$$\begin{aligned} \Gamma_{1,CD} &= \Gamma_{CD}, \\ \Gamma_{2,CD} &= \Gamma_{CD} + I_C \oplus 0_D, \\ \Gamma_{3,CD} &= \Gamma_{CD} + 0_C \oplus I_D, \\ \Gamma_{4,CD} &= \Gamma_{CD} + I_{CD}. \end{aligned} \quad (4.23)$$

Correlation Coefficient $E(\theta_i, \phi_k)$

The joint probability distribution $P(x_{\theta_i}^A, x_{\phi_k}^B)$ of the quadratures $x_{\theta_i}^A$ and $x_{\phi_k}^B$ appearing in the formula (4.5) for the correlation coefficient $E(\theta_i, \phi_k)$ can be obtained from the Wigner function (4.19) as a marginal distribution. We have

$$P(x_{\theta_i}^A, x_{\phi_k}^B) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W_{c,AB}(S_{\text{sh}}^T r_{\theta_i, \phi_k}) dp_{\theta_i}^A dp_{\phi_k}^B, \quad (4.24)$$

where $r_{\theta_i, \phi_k} = [x_{\theta_i}^A, p_{\theta_i}^A, x_{\phi_k}^B, p_{\phi_k}^B]$ and the symplectic matrix $S_{\text{sh}} = S_{\text{PS},A}(\theta_i) \oplus S_{\text{PS},B}(\phi_k)$ describes local phase shifts applied to modes A and B that map the measured quadratures $x_{\theta_i}^A$ and $x_{\phi_k}^B$ onto the quadratures x^A and x^B , respectively.

In order to express the result of the integration in Eq. (4.24) in a compact matrix notation, we re-order the elements of the vector r_{θ_i, ϕ_k} as follows,

$$\begin{bmatrix} x_{\theta_i}^A \\ x_{\phi_k}^B \\ p_{\theta_i}^A \\ p_{\phi_k}^B \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{\theta_i}^A \\ p_{\theta_i}^A \\ x_{\phi_k}^B \\ p_{\phi_k}^B \end{bmatrix} \quad (4.25)$$

which defines a matrix S_{hom} . After these algebraic manipulations, the four matrices $\Gamma_{j,AB}$ appearing in the exponents in Eq. (4.19) transform to

$$\Gamma'_{j,AB} = S_{\text{hom}} S_{\text{sh}} \Gamma_{j,AB} S_{\text{sh}}^T S_{\text{hom}}^T \equiv \begin{bmatrix} A_j & C_j \\ C_j^T & B_j \end{bmatrix}, \quad (4.26)$$

where we have divided the matrix $\Gamma'_{j,AB}$ into four sub-matrices with respect to the x vs p splitting. A straightforward integration over $p_{\theta_i}^A$ and $p_{\phi_k}^B$ in Eq. (4.24) then yields the joint probability distribution,

$$P(x_{\theta_i}^A, x_{\phi_k}^B) = \frac{\sqrt{\det \Gamma_{\text{out}}}}{\pi P_G} \sum_{j=1}^4 \frac{q_j e^{-y^T \Gamma_j y}}{\sqrt{\det \Gamma_{j,CD}} \sqrt{\det B_j}}, \quad (4.27)$$

where $y = (x_{\theta_i}^A, x_{\phi_k}^B)^T$ and

$$\Gamma_j = A_j - C_j B_j^{-1} C_j^T. \quad (4.28)$$

Taking into account the choice of binning, the normalization of the joint probability distribution, and its symmetry, $P(x_{\theta_i}^A, x_{\phi_k}^B) = P(-x_{\theta_i}^A, -x_{\phi_k}^B)$, we can express the correlation coefficient as follows,

$$E(\theta_i, \phi_k) = 4 \int_0^\infty \int_0^\infty P(x_{\theta_i}^A, x_{\phi_k}^B) dx_{\theta_i}^A dx_{\phi_k}^B - 1. \quad (4.29)$$

This last integral can be easily evaluated analytically. For a given Γ_j matrix

$$\Gamma_j = \begin{bmatrix} a_j & c_j \\ c_j & b_j \end{bmatrix}, \quad (4.30)$$

the integral of the exponential term

$$G_j = \int_0^\infty \int_0^\infty e^{-a_j y_1^2 - b_j y_2^2 - 2c_j y_1 y_2} dy_1 dy_2 \quad (4.31)$$

can be calculated by transforming to polar coordinates and integrating first over the radial coordinate and then over the angle. After some algebra (see Appendix G), we finally arrive at

$$G_j = \frac{1}{2\sqrt{a_j b_j - c_j^2}} \left[\frac{\pi}{2} - \arctan \frac{c_j}{\sqrt{a_j b_j - c_j^2}} \right]. \quad (4.32)$$

The final fully analytical formula for the correlation coefficient reads

$$E(\theta_i, \phi_k) = \frac{4\sqrt{\det \Gamma_{\text{out}}}}{\pi P_G} \left[\sum_{j=1}^4 \frac{q_j G_j}{\sqrt{\det \Gamma_{j,CD}} \sqrt{\det B_j}} \right] - 1 \quad (4.33)$$

and the Bell factor can be expressed as

$$S = E(\theta_1, \phi_1) + E(\theta_1, \phi_2) + E(\theta_2, \phi_1) - E(\theta_2, \phi_2). \quad (4.34)$$

Violation of Bell-CHSH Inequalities

A necessary condition for the observation of a violation of Bell inequalities with homodyne detectors is that the Wigner function of the two-mode state used in the Bell test is not positive definite. Figure 4.7 illustrates that the Wigner function (4.19) of the conditionally generated state $\rho_{c,AB}$ is indeed negative in some regions of the phase space. The area of negativity, as well as the attained negative values of W , are rather small, which indicates that we should not expect a high Bell violation with homodyning.

As we have shown before, the maximum Bell factor S achievable with our setup and sign binning is about $S = 2.048$. We conjecture that this binning is optimal or close to optimal. This is supported by the simple structure of the joint probability distribution

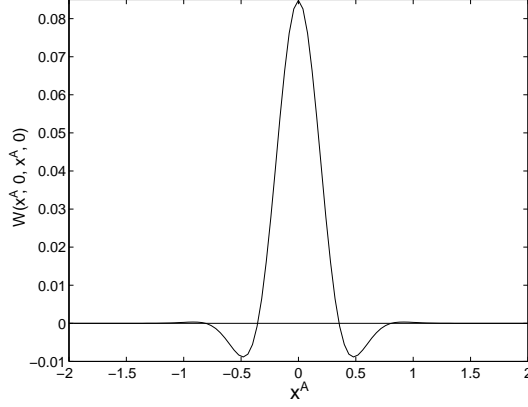


Figure 4.7: A one-dimensional cut of the Wigner function of the two-mode state $\rho_{c,AB}$ along the line $x^B = x^A$, $p^A = p^B = 0$ for $\lambda = 0.6$ and beam splitters BS_A and BS_B transmittance $T = 0.95$. Notice the regions where W is negative.

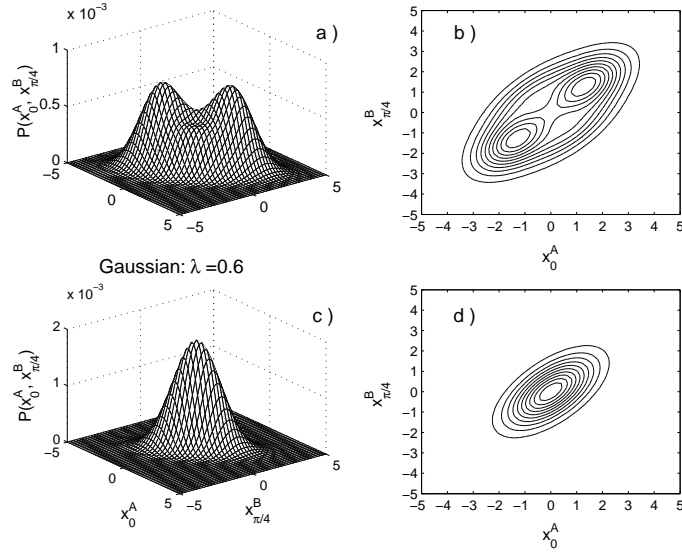


Figure 4.8: Joint probability distribution $P(x_{\theta_j}^A, x_{\phi_k}^B)$. Panels (a) and (b) show the distribution for the conditionally-prepared non-Gaussian state with $T = 0.99$. Panels (c) and (d) display the distribution for the initial Gaussian two-mode squeezed vacuum state. The curves are plotted for perfect detectors $\eta_{PD} = \eta_{BHD} = 100\%$, squeezing $\lambda = 0.6$ and $\theta_{\text{Alice}} = 0$ and $\phi_{\text{Bob}} = \pi/4$.

(4.27). As can be seen in Fig. 4.8(a,b), P exhibits two peaks, both located in the quadrants where Alice's and Bob's measured quadratures have the same sign. Note also that the two-peak structure is a clear signature of the non-Gaussian character of the state (c.f. Fig. 4.8(c,d)).

We have carried out numerical calculations of S for several other possible binning which divide the quadrature axis into three or four intervals, and have not found any binning which would provide higher S than the sign binning. We have also performed optimization over the angles θ_j and ϕ_k and all the results and figures presented in this

Section were obtained for the optimal choice of angles $\theta_1 = 0$, $\theta_2 = \pi/2$, $\phi_1 = -\pi/4$, $\phi_2 = \pi/4$.

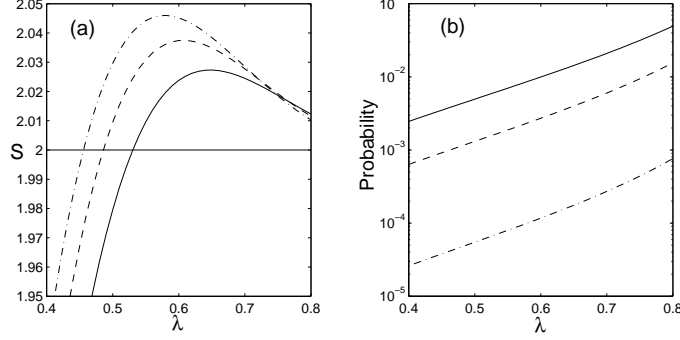


Figure 4.9: Violation of Bell-CHSH inequality with the conditionally-prepared non-Gaussian state. (a) Bell factor S as a function of the squeezing. (b) Probability of success of the generation of the non-Gaussian state as a function of the squeezing. The curves are plotted for perfect detectors ($\eta_{PD} = \eta_{BHD} = 100\%$) with $T = 0.9$ (solid line), $T = 0.95$ (dashed line), and $T = 0.99$ (dot-dashed line).

Figure 4.9(a) illustrates that the Bell-CHSH inequality $|S| \leq 2$ can be violated with the proposed set-up, and shows that there is an optimal squeezing λ_{opt} which maximizes S . This optimal squeezing is well predicted by the simple model assuming perfect detectors with single-photon resolution, $\lambda_{\text{opt}}T \approx 0.57$. The curve plotted for $T = 0.99$ practically coincides with the results obtained from the ideal model, c.f. Fig. 4.5(a). This confirms that in the limit $T \rightarrow 1$ the detectors with single-photon sensitivity become for our purposes equivalent to photodetectors with single-photon resolution. The maximum Bell factor achievable with our scheme is about $S_{\text{max}} \approx 2.045$ which represents a violation of the Bell inequality by 2.2%. To get close to the S_{max} one needs sufficiently high (but not too strong) squeezing. In particular, the value $\lambda \approx 0.57$ corresponds to approximately 5.6 dB of squeezing. Figure 4.9(b) illustrates that there is a clear trade-off between S and the probability of success P_G . To maximize S one should use highly transmitting beam splitters but this would reduce P_G . The optimal T that should be chosen would clearly depend on the details of the experimental implementation.

Sensitivity to the Experimental Imperfections

It is shown in Fig. 4.10(a) that the Bell factor S depends only very weakly on the efficiency η_{PD} of the single-photon detectors, so the Bell inequality can be violated even if $\eta_{PD} \approx 1\%$. This is very important from the experimental point of view because, although the quantum detection efficiencies of the avalanche photodiodes may be of the order of 50%, the necessary spectral and spatial filtering which selects the mode that is detected by the photodetector may reduce the overall detection efficiency to a few percent. Low detection efficiency only decreases the probability of conditional generation P_G of the non-Gaussian state, see Fig. 4.10(b). The dependence of P_G on η_{PD} and T can be very well approximated by a quadratic function, $P_G \approx \eta_{PD}^2(1 - T)^2$ which quickly drops when η_{PD} decreases. In practice, the minimum necessary η_{PD} will be determined mainly by the constraints on the total time of the experiment and by the dark counts of the detectors.

In contrast, the Bell factor S strongly depends on the efficiency of the homodyne detectors, and η_{BHD} must be above $\sim 90\%$ in order to observe Bell violation, see

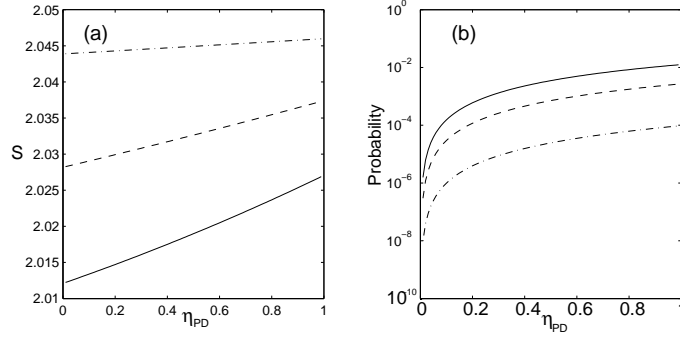


Figure 4.10: Effect of the inefficiency of the photodetectors PD_A and PD_B . (a) Bell parameter S as a function of the efficiency η_{PD} of the photodetectors. (b) Probability of success as a function of the efficiency η_{PD} . The curves are plotted for $T\lambda = 0.57$, $\eta_{BHD} = 100\%$ and the same transmittance as in Fig. 4.9.

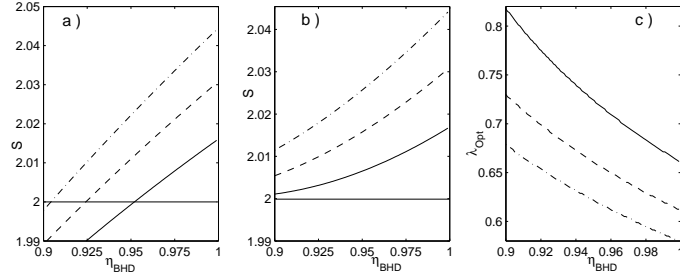


Figure 4.11: Effect of inefficient homodyning. (a) Bell parameter S as a function of the efficiency η_{BHD} of the homodyning. The curve is plotted for $T\lambda = 0.57$, $\eta_{PD} = 30\%$ and the same transmittance as in Fig. 4.9. (b) Bell parameter achieved for the optimal squeezing λ_{opt} is plotted as a function of η_{BHD} . (c) Optimal squeezing λ_{opt} is plotted as a function of η_{BHD} . The curve is plotted for $\eta_{PD} = 30\%$ and the same transmittance as in Fig. 4.9.

Fig. 4.11. However, this is not an obstacle because such (and even higher) efficiency has been already achieved experimentally (see *e.g.* [203]). Interestingly, we have found that it is possible to partially compensate for imperfect homodyning with efficiency $\eta_{BHD} < 1$ by increasing the squeezing of the initial state. This effect is illustrated in Fig. 4.11(b) which shows the dependence of the Bell factor S on η_{BHD} for optimal squeezing λ_{opt} . Figure 4.11(c) then shows how the optimal squeezing increases with decreasing η_{BHD} .

In addition to imperfect detection efficiency η_{BHD} , the electronic noise of the homodyne detector is another factor that may reduce the observed Bell violation. We model the added electronic noise by assuming that the effective quadrature that is detected \hat{x}_{det} is related to the signal quadrature \hat{x}_S by a formula,

$$\hat{x}_{det} = \sqrt{\eta_{BHD}} \hat{x}_S + \sqrt{1 - \eta_{BHD}} \hat{x}_{vac} + \sqrt{N_{el}} \hat{x}_{noise},$$

where \hat{x}_{vac} and \hat{x}_{noise} are two independent Gaussian distributed quadratures with zero mean and variance $1/2$, and N_{el} is the electronic noise variance expressed in shot noise units. On the level of covariance matrices, N_{el} can be included by modifying the

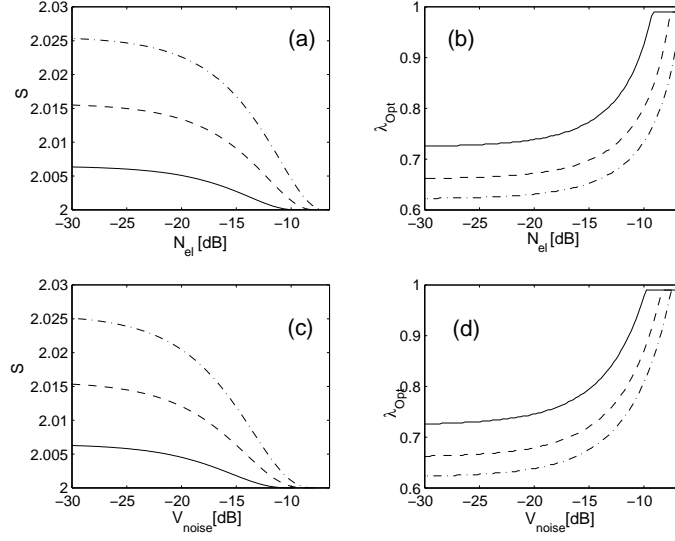


Figure 4.12: Effect of the electronic noise and thermal input states. (a) Maximum achievable Bell parameter S with the optimal squeezing λ_{opt} as a function of the electronic noise N_{el} . (b) Optimal squeezing λ_{opt} giving the highest Bell parameter S for a given electronic noise. (c) Maximum Bell parameter S as a function of the thermal noise of the input state V_{noise} . (d) Optimal squeezing λ_{opt} giving the highest Bell parameter S for a given thermal noise at the input. The curves are plotted for $\eta_{\text{PD}} = 30\%$, $\eta_{\text{BHD}} = 95\%$, and $T = 0.9$ (solid line), $T = 0.95$ (dashed line), and $T = 0.99$ (dot-dashed line).

formula for the noise matrix G ,

$$G = (1 - \eta_{\text{BHD}} + N_{\text{el}})I_{AB} \oplus (1 - \eta_{\text{PD}})I_{CD}. \quad (4.35)$$

The homodyne detector with electronic noise is actually equivalent to a detector without noise but with a lower homodyne detector efficiency $\eta'_{\text{BHD}} = \eta_{\text{BHD}}/(1 + N_{\text{el}})$. This can be shown by noting that the re-normalized quadrature $x_{\text{det}}/\sqrt{1 + N_{\text{el}}}$ is exactly a quadrature that would be detected by a balanced homodyne detector with $N_{\text{el}} = 0$ and efficiency η'_{BHD} . Our calculations reveal that the electronic noise should be 15–20 dB below shot noise (see Fig. 4.12(a) and (b)), which is currently attainable with low-noise charge amplifiers. Again, higher squeezing can partially compensate for the increasing noise.

So far we have assumed that the source in Fig. 1 emits pure two-mode squeezed vacuum state. However, experimentally, it is very difficult to generate pure squeezed vacuum saturating the Heisenberg inequality. It is more realistic to consider a mixed Gaussian state such as squeezed thermal state which can be equivalently represented by adding quadrature independent Gaussian noise with variance V_{noise} to each mode of the two-mode squeezed vacuum. The effect of the added noise stemming from input mixed Gaussian state is quite similar to the influence of the electronic noise of the homodyne detector, see Fig. 4.12 (c) and (d). We find again that the added noise in the initial Gaussian state should be 15–20 dB below the shot noise.

In the experimental demonstration of single-photon subtraction [197], a main source of noise and imperfections was that the single-photon detector was sometimes triggered by a photon coming from other modes than the mode detected in the balanced homodyne detector. The single-mode description of a parametric amplifier is only an ap-

proximation and the amplifier produces squeezed vacuum in several modes. A balanced homodyne detector very efficiently selects a single mode defined by the spatiotemporal profile of the local oscillator pulse. However, this reference is missing in case of single-photon detector, where the effective single mode has to be selected by spatial and spectral filtering, which reduces the overall detection efficiency η . In practice, the filtering is never perfect, hence the photodetector PD_A (PD_B) can sometimes click although no photon was removed from mode A (B).

We can model this false triggering by re-defining the POVM element $\Pi_{1,C}$ ($\Pi_{1,D}$) appearing in Eq. (4.17). The new Π_1 becomes a convex mixture of the original POVM element $I - |0\rangle\langle 0|$, which corresponds to triggering by a photon coming from the mode A(B), and the identity operator I , which corresponds to the false triggering. We can write $\Pi_1(\xi) = I - \xi|0\rangle\langle 0|$ and the coefficient $0 \leq \xi \leq 1$ can be related to the fraction of false triggers P_f . Assuming for simplicity pure two-mode squeezed vacuum in modes A and B, the single-mode state in C or D just before detection is a thermal state with mean number of chaotic photons $\bar{n} = \eta_{\text{PD}}(1 - T)\lambda^2/(1 - \lambda^2)$ (note that this includes the effect of imperfect detectors with efficiency η_{PD}). The probability of projection of the thermal state on vacuum reads $P_{\text{vac}} = 1/(\bar{n} + 1)$. The probability of false trigger P_f can be expressed in terms of the probability of a trigger $P(\xi) = 1 - \xi P_{\text{vac}}$ and the probability of a correct triggering event $P(\xi = 1) = 1 - P_{\text{vac}}$,

$$P_f = \frac{P(\xi) - P(\xi = 1)}{P(\xi)}. \quad (4.36)$$

From this formula we obtain

$$\xi = \frac{1 - (1 + \bar{n})P_f}{1 - P_f}. \quad (4.37)$$

The analytical formula (4.34) for the Bell factor S can still be used even in the presence of false triggering. We only have to re-define the four coefficients q_j as follows, $q_1 = 1$, $q_2 = q_3 = -2\xi$, $q_4 = 4\xi^2$.

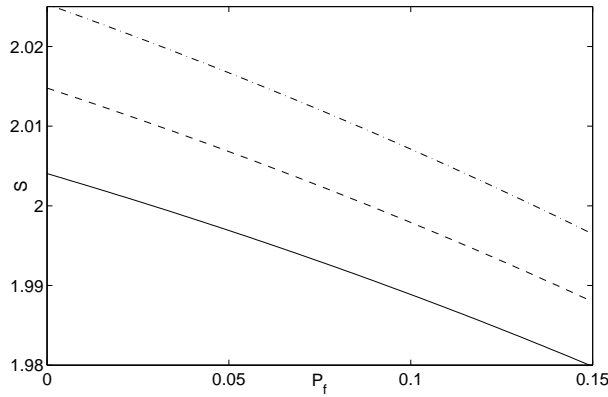


Figure 4.13: Influence of false triggers. The Bell factor S is plotted as a function of the probability of false triggering P_f for $T = 0.9$, $\lambda = 0.62$ (solid line), $T = 0.95$, $\lambda = 0.66$ (dashed line), and $T = 0.99$, $\lambda = 0.72$ (dot-dashed line), $\eta_{\text{PD}} = 30\%$, and $\eta_{\text{BHD}} = 95\%$.

The effect of the false triggers is illustrated in Fig. 4.13. As expected, the achievable Bell factor decreases with increasing P_f . The results are shown for a realistic set of parameters as identified in [85] and for three different values of T . For high transmittance ($T = 0.99$) up to 11% of false triggers can be tolerated while for $T = 0.95$ the acceptable fraction of false triggers decreases to $P_f = 6\%$. In a recent experiment

[197], the estimated fraction of false triggers was $P_f \approx 30\%$ which would have to be significantly reduced in the Bell test experiment. Possible ways of suppressing false triggers include better filtering and/or using sources that produce squeezed light in well defined spatial modes, such as nonlinear periodically poled waveguides.

Four Photon Subtractions

Until now we have focused on a single-photon subtraction on each side (one photon removed from mode A and one from mode B). If we now consider a scheme where two photons are subtracted from each mode, the de-Gaussification of the state will be stronger and we may expect a higher Bell violation than before. To subtract two photons from each mode, we only need to add one more unbalanced beam splitter and photodetector on each side in Fig. 4.6. A successful state generation would be indicated by simultaneous clicks of all four detectors. Assuming perfect photon-number resolving detectors, the state generated from two-mode squeezed vacuum (4.6) by subtracting two photons from each mode can be expressed as (see Appendix F),

$$\begin{aligned} |\psi_{\text{out}}\rangle_{AB} &\propto \hat{a}_A^2 \hat{a}_B^2 |\psi_{\text{in}}(T^2\lambda)\rangle_{AB} \\ &\propto \sum_n (n+2)(n+1)(T^2\lambda)^n |n, n\rangle_{AB}, \end{aligned} \quad (4.38)$$

and the probability of success reads

$$\mathcal{P}_{4\text{ph}} = 4T^2(1-T)^4\lambda^4(1-\lambda^2) \frac{1 + 4T^4\lambda^2 + T^8\lambda^4}{(1 - T^4\lambda^2)^5}. \quad (4.39)$$

Since the state (4.38) exhibits perfect photon number correlations, the Munro's formula for the Bell factor can again be directly applied [134]. Numerical calculations show that the maximum Bell violation with the state (4.38) and sign binning of quadratures is achieved for $T^2\lambda = 0.40$ which yields $S_{\text{max},4\text{ph}} = 2.064$, which is indeed higher than the maximum achievable with two-photon subtraction, $S_{\text{max},2\text{ph}} = 2.048$, and very close to the maximum value $S = 2.076$ [134].

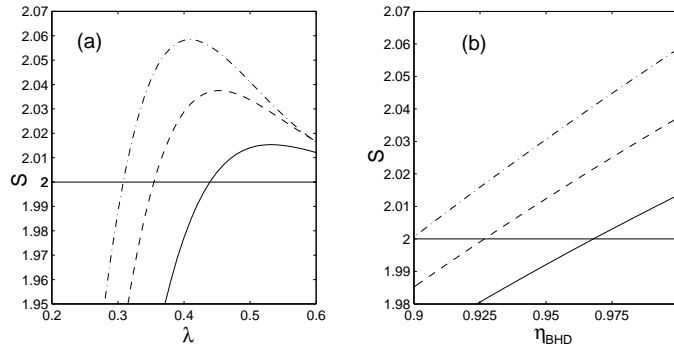


Figure 4.14: Violation of Bell-CHSH inequality with four photon subtractions. (a) Bell parameter S as a function of the squeezing λ for perfect detectors $\eta_{\text{PD}} = \eta_{\text{BHD}} = 100\%$. (b) Bell parameter S as a function of the efficiency η_{BHD} of the homodyning. The curve is plotted for $T^2\lambda = 0.40$, $\eta_{\text{PD}} = 100\%$ and the same transmittance as in Fig. 4.9.

A more realistic description of the four-photon subtraction scheme that takes into account realistic imperfect photon subtraction can be developed using the approach used for two photon subtractions. We find that the Wigner function of the conditionally

generated state is a linear combination of sixteen Gaussian. The results of numerical calculations are shown in Figs. 4.14(a) and (b), which illustrate that the two-photon subtraction from each mode yields higher violation of Bell-CHSH inequality than one-photon subtraction only for very high transmittance $T > 0.95$. For lower transmittance, the fact that the photodetectors do not distinguish the number of photons reduces the Bell factor. Moreover, adding a second stage of photon subtractions dramatically decreases the probability of generating the non-Gaussian state. The probability can be estimated as $P_G \approx \eta_{PD}^4 (1-T)^4$, so for $T > 0.95$ and $\eta = 50\%$ we get $P_G \approx 10^{-6}$ and the duration of data acquisition would make the experiment infeasible. We conclude that from the practical point of view there seems to be no advantage in using the scheme with four photon subtractions instead of the much simpler scheme with two photon subtractions.

4.5 Alternative Schemes

In this section we will study the violation of Bell-CHSH inequalities for a large group of alternative schemes, which involve from one to four photon subtractions. The main objective of this section is to compare the maximum Bell-CHSH factor S obtained for the different proposed setups. As the main purpose of this section is the comparison of the different schemes, we will consider only idealized schemes with almost perfect single-photon subtraction on the beam splitters ($T = 0.99$), and perfect photodetectors and homodyning ($\eta_{PD} = \eta_{BHD} = 100\%$). The maximum achievable Bell factor for each scheme presented below was determined by optimizing over the angles $\theta_{1,2}$, $\phi_{1,2}$ as well as over the squeezing λ of the initial Gaussian states. The sign binning of the measured quadratures has been used in all cases. All the schemes presented in this section use the symbol convention depicted in Fig. 4.15.

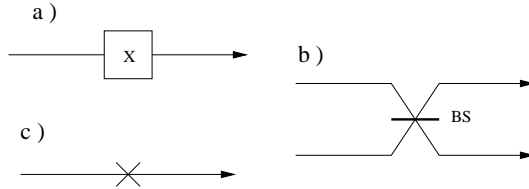


Figure 4.15: Symbol convention. (a) Single-mode squeezer along the x quadrature. (b) Beam splitter. (c) Conditional subtraction of a photon as described in the preceding section.

In the preceding section, we have seen that the probability of successful generation of a non-Gaussian state decreases significantly with the number of photon subtractions. At the same time the complexity of the implementation of the experimental setup increases with the number of photon subtractions. It is then obvious that the most interesting schemes for a Bell-CHSH violation are those involving only one photon subtraction. Unfortunately, for the schemes that we have considered (see Fig. 4.16), no violation was observed¹. In this case, the maximal value of the Bell-CHSH factor is $S = 2$, which is achieved at the limit of an infinite squeezing. Note that in this limit the effect of subtracting a single photon vanishes, so that one clearly tends to an infinitely squeezed Gaussian state $\sum_n |n\rangle_A |n\rangle_B$. For this state, it is easy to check that

¹Note that we represent the two-mode squeezer using its theoretical equivalent scheme composed of two orthogonal single-mode squeezers followed by a beam splitter. Even though these two schemes correspond to physically distinct optical implementations, this choice of representation is better adapted to the comparison between the different possible positions of the photon subtraction.

$S = 2$ can be achieved with the sign binning and an appropriate choice of angles. So, here and below, all the schemes that do not result in a Bell violation correspond to $S = 2$, a point which is associated with the limit $\lambda \rightarrow 1$.

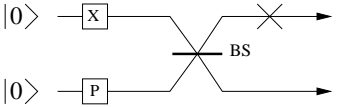
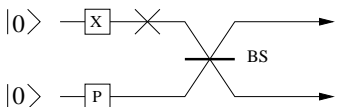
	Schemes: one subtraction	S
a)		2
b)		2

Figure 4.16: Schemes with only one photon subtraction. The first column labels the different setups proposed, the second shows the scheme and finally the last column gives the maximal Bell factor S obtained when optimizing the squeezing. (a) Photon subtraction after the creation of the two-mode squeezed vacuum. (b) Photon subtraction before mixing two single-mode squeezed states on a beam splitter.

After one photon subtraction, the simplest schemes are those with two photon subtractions. In the preceding sections it was shown that it is possible to violate the Bell-CHSH inequality with two photon subtractions (scheme Fig. 4.17(a)). It follows from Fig. 4.17 that several other schemes (see Figs. 4.17(d) and (e)) also violate Bell-CHSH inequality, but the maximal achievable Bell factor S appears to be much smaller in comparison to the scheme shown in Fig. 4.17(a).

By adding one more photon subtraction to the schemes shown in Fig. 4.17, we can construct an ensemble of schemes with three photon subtractions. After numerical optimization we have found that none of these schemes succeeds to violate Bell-CHSH inequality. This striking result together with the fact that we have not found any violation for schemes based on a single subtraction suggests that it may be necessary to have a scheme with an even number of photon subtractions in order to observe $S > 2$.

In the preceding section, we have also proposed one scheme with four photon subtractions that violates Bell-CHSH inequality. Many other possible schemes exist where four photons are subtracted. Figure 4.18 illustrates some particular examples, which are based on the preparation of two-mode squeezed vacuum via mixing of two single-mode squeezed states on an balanced beam splitter. The photon subtractions are symmetrically placed to both modes. Strikingly, if all four photons are subtracted either before or after mixing on a beam splitter, then we get $S > 2$. However, if a single photon is subtracted from each mode both before and after combining the modes on a beam splitter, then we do not obtain any Bell violation.

Finally we have also studied an alternative group of schemes where instead of subtracting photons separately from modes A and B , we mix the auxiliary modes C and D on a balanced beam splitter before the detection on the photodetectors. Consider the scheme depicted in Fig. 4.19(a) where only a single photon is subtracted. The mixing of modes C and D on a beam splitter erases the information about the origin of the detected photon which implies that the conditionally prepared state is a coherent superposition of states where a single photon has been removed either from mode A or from mode B . However, even this modification does not lead to Bell violation with

	Schemes: two subtractions	S
a)		2.046
b)		2
c)		2
d)		2.02
e)		2.01

Figure 4.17: Schemes with two photon subtractions. The right column gives the maximal value of the Bell factor S for the proposed setups.

	Schemes: four subtractions	S
a)		2.06
b)		2.05
c)		2

Figure 4.18: Schemes with four photon subtractions. Last column gives the maximal value of the Bell factor S for the proposed setups.

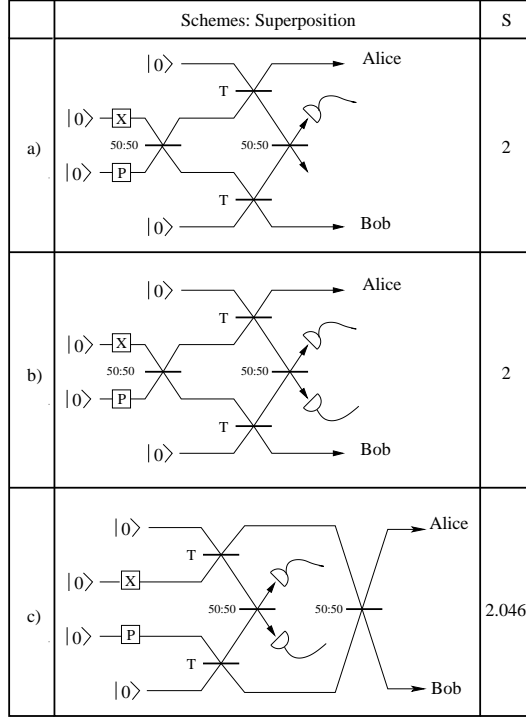


Figure 4.19: Schemes consisting of superpositions of other schemes proposed above. (a) Superposition of one photon subtraction on modes A or B . (b), (c) Superposition of two photon subtractions on modes A or B .

just a single subtraction.

We can extend the scheme by placing a photodetector at both output ports of the beam splitter, cf. Fig. 4.19(b). In the limit of a high transmittance $T \rightarrow 1$, the conditioning on the click of each detector selects the events where there were altogether two photons at the beam-splitter inputs. The bosonic properties of the photons imply that a simultaneous click of both photodetectors occurs only if the two subtracted photons are coming from the same mode (A or B) [106], but again we do not know from which mode the two photons are subtracted. This scheme is thus equivalent to the superposition of two schemes of the type shown in Fig. 4.17(c). Unlike the scheme in Fig. 4.17(c), the scheme in Fig. 4.19(b) is symmetric with respect to the modes A and B . However, no violation can be observed. On the other hand, the scheme in Fig. 4.19(c) leads to $S > 2$ by realizing a superposition of states where two photons are subtracted from a single-mode squeezed vacuum state and this state is then mixed with another single-mode squeezed vacuum on a balanced beam splitter, see 4.17(d). In comparison to the scheme in Fig. 4.17(d), we obtain much higher violation ($S = 2.046$), recovering the optimal for two photon subtractions.

4.6 Setup Proposal and Realistic Parameters

In our experimental proposal, Fig. 4.20, the source (controlled by Sophie) is based on a master laser beam, which is converted into second harmonic in a nonlinear crystal (SHG). After spectral filtering (F), the second harmonic beam pumps an optical parametric amplifier (OPA) which generates two-mode squeezed vacuum in modes A and

B. Single photons are conditionally subtracted from modes A and B with the use of the beam splitters BS_A and BS_B and single-photon detectors PD_A and PD_B . Alice (Bob) measures a quadrature of mode A (B) using a balanced homodyne detector that consists of a balanced beam splitter BS_3 (BS_4) and a pair of highly-efficient photodiodes. The local oscillators LO_A and LO_B are extracted from the laser beam by means of two additional beam splitters BS_1 and BS_2 . The random switching of the relative phase θ (ϕ) between LO_A and A (LO_B and B) can be performed using fast electro-optical modulators.

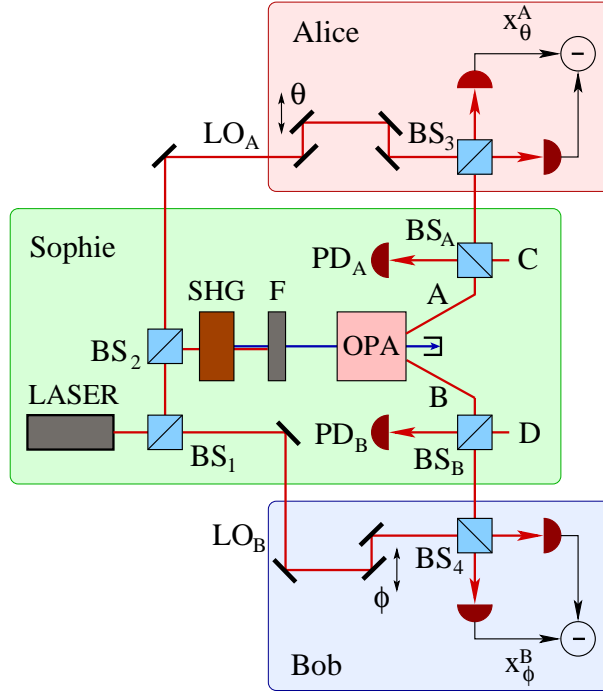


Figure 4.20: Proposed experimental setup.

In order to be more specific, let us consider the single-mode photon subtraction experiment [197]. It is based on a commercial cavity-dumped titanium-sapphire laser, delivering nearly Fourier-limited pulses at 850 nm, with a duration of 150 fs and a repetition rate of 790 kHz. Squeezed vacuum pulses generated by parametric deamplification are sent through a beam splitter, and the reflected beam is detected by a silicon APD. Conditional on a click, the transmitted pulse is prepared in a non-Gaussian state, which is measured by homodyne detection with an overall efficiency $\eta_{\text{BHD}} \approx 75\%$. This experiment gives us useful estimates for a possible Bell test. First, the delay between pulses ($1.2 \mu\text{s}$) allows ample time for individual pulse analysis. A fast random choice of the analyzed quadratures can be performed using electro-optical modulators on the LO beams, triggered for instance by digitizing the shot-noise of locally generated auxiliary beams. Switching times around 100 ns, associated with propagation distances of a few tens of meters, seems quite feasible. The APDs can be triggered only when a pulse is expected, reducing the effect of dark counts to a negligible value. The intrinsic APD efficiency is about 50%, but the filtering used to select a single mode currently reduces the overall η to less than 5%, which should be improved for accumulating enough statistics.

This allows us to define a set of realistic parameter values that should be reached in a loophole-free Bell test : with $\eta = 30\%$, $T = 95\%$, and $\lambda = 0.6$, a violation of

Bell-CHSH inequality by about 1% should be observable if the homodyne efficiency η_{BHD} is larger than 95%. With a repetition rate of 1 MHz and $P \approx 2.6 \times 10^{-4}$, the number of data samples would be several hundreds per second, so that the required statistics to see a violation in the percent range could be obtained in a reasonable time (a few hours). In addition, the electronic noise of the homodyne detectors should be 15-20 dB below shot noise, attainable with low-noise charge amplifiers. All these numbers have already been reached separately in various experiments, but attaining them simultaneously certainly represents a serious challenge. Nevertheless, taking into account many possible experimental improvements, the existence of an experimental window for a loophole-free test of Bell inequalities can be considered as highly plausible. Therefore, it appears that, with quantum continuous variables, a reasonable compromise can be found between the experimental constraints and the very stringent requirements imposed by a loophole-free test of Bell inequalities.

4.7 Conclusions

We have proposed an experimentally feasible setup allowing for a loophole-free Bell test with efficient homodyne detection using a non-gaussian entangled state generated from a two-mode squeezed vacuum state by subtracting a single photon from each mode. We have presented a full analytical description of a realistic setup with imperfect detectors, noise and mixed input states. We have studied in detail the influence of the detector inefficiencies, the electronic noise of homodyne detector, and the input mixed states, on the achievable Bell violation. The main feature of the present scheme is that it is largely insensitive to the detection efficiency of the avalanche photodiodes that are used for conditional preparation of the non-gaussian state, so that detector efficiencies of the order of a few per cent are sufficient. On the other hand, the detection efficiency of the balanced homodyne detector should be of the order of 90% and the electronic noise of the homodyne detector should be at least 15 dB below the shot noise level. The optimal squeezing that yields maximum Bell violation depends on the experimental circumstances but is, generally speaking, within the range of experimentally attainable values. As a rule, the optimal squeezing increases with decreasing η_{BHD} and increasing noise.

We have also discussed several alternative schemes that involve the subtraction of one, two, three or four photons. Unfortunately, the experimentally simplest and most appealing scheme consisting in a single photon subtraction does not exhibit violation of the proposed Bell inequalities. Taking into account that we have not found any scheme with three photon subtractions which would violate Bell-CHSH inequality, the only way of exceeding the 2.046 violation appears to be by subtracting four photons. Unfortunately, the price to pay for this slight increase of S is that the probability of successful conditional generation is so low that it makes the experiment infeasible.

The experimental demonstration of a single photon subtraction from a single-mode squeezed vacuum state [142, 197] provides a strong incentive for further theoretical and experimental developments along these lines, and we can thus expect that some of the schemes discussed here will be experimentally implemented in a not too distant future.

Part II

Continuous-Variable Quantum Key Distribution

Chapter 5

Shannon Information Theory

5.1 Introduction to Data Compression

Let us consider an experiment where we toss a fair coin many times. Even if it is impossible to guess the outcome of a given toss, we expect after a large number of events that half of the results are head. Now repeat the experiment with a biased coin with probability of outputting head $p_h = p$. The law of large numbers [133] tells us that after a large number (n) of coin tossing the amount of head (N_h) and tails (N_t) will be approximately, $N_h \approx np$ and $N_t \approx n(1-p)$, respectively. All the sequences that we observe after a high number of coin tossing belong to a subset called the *typical set* that asymptotically captures all the occurrence probability.

Typical Sequence

A sequence of n coin tossing is called *typical* if it is composed of $N_h \approx np$ heads and $N_t \approx n(1-p)$ tails, its occurrence probability reads

$$p(x_1, x_2, \dots, x_n) \approx p^{np}(1-p)^{n(1-p)}, \quad (5.1)$$

which gives

$$\log p(x_1, x_2, \dots, x_n) \approx np \log p + n(1-p) \log(1-p) = -nH(p), \quad (5.2)$$

where $H(p)$ is the Shannon entropy,

$$H(p) = -[p \log p + (1-p) \log(1-p)], \quad (5.3)$$

where the log is a base 2 logarithm and the entropy is expressed in bits. Remark that, by convention, $0 \log 0 = 0$ which is justified by continuity. This function displayed in Fig. 5.1, is concave with minima ($H(0) = H(1) = 0$) corresponding to deterministic coins, and maximum for fair random coins $H(1/2) = 1$.

By definition all the typical sequences have the same probability of occurrence $p_{TS} \approx 2^{-nH(p)}$. The number of such typical sequences N_{TS} can be estimated for larger n assuming that np is (close to) an integer, as

$$N_{TS} \approx \binom{n}{np} = \frac{n!}{(np)!(n(1-p))!}. \quad (5.4)$$

Using the Stirling formula for large numbers

$$\log n! \approx n \log n - n, \quad (5.5)$$

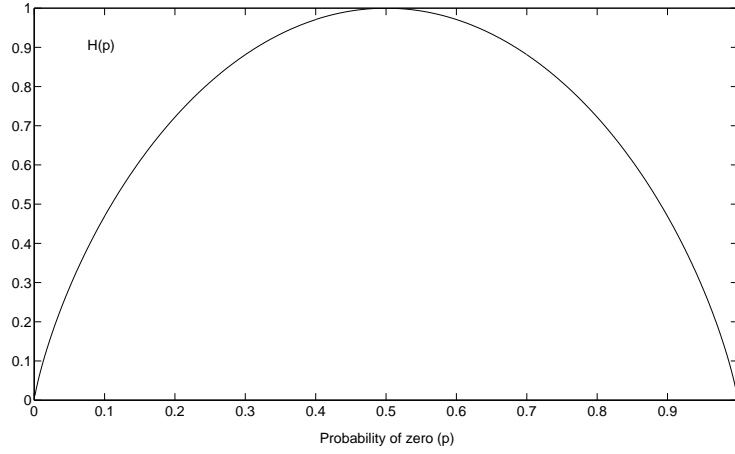


Figure 5.1: Entropy of a binary source, where p is the probability of one of the outcomes.

one can show that the number of typical sequences is approximately

$$N_{TS} \approx 2^{nH(p)}. \quad (5.6)$$

We conclude that among the set of all possible sequences that may occur after tossing n coins ($N = 2^n$), there is a subset composed of approximately $N_{TS} \approx 2^{nH(p)}$ equiprobable typical sequences ($p_{TS} \approx 2^{-nH(p)}$) that concentrates almost all the occurrence probability, $P_{TS} = p_{TS}N_{TS} \approx 1$.

Data Compression

Imagine that you want to send by email the result of the n coin tossing $x^n = x_1, x_2, \dots, x_n$. A trivial solution will be to send a string of n bits where heads are encoded by zeros and tails by ones. But, does there exist a more economic technique that allows one to reduce the number of bits we have to send. The answer is yes, for example using the following encoding-decoding (see Fig. 5.2) based on the properties of the typical set:

1. The encoding operation of x^n :
 - If x^n is typical, use a bijective mapping \mathcal{M} that encode which of the $2^{nH(p)}$ typical sequences you have obtained using $nH(p)$ bits.
 - If its not typical send a string of $nH(p)$ zeros.
2. The decoding of x^n :
 - Apply the inverse mapping \mathcal{M}^{-1} in order to recover the typical sequences from the $nH(p)$ bits.

We observe that an error occurs only when the sequence x^n is non-typical. The probability of the typical set being arbitrarily close to one for $n \rightarrow \infty$ the error can be made arbitrarily small, implying that the encoding is reliable. Later in this chapter we will show that $nH(p)$ is indeed the optimal compression that can be achieved.

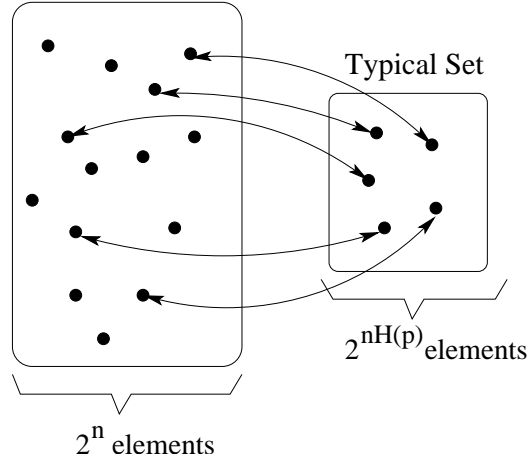


Figure 5.2: The encoding bijective function \mathcal{M} maps the typical set into an smaller ensemble of $nH(X)$ bits. The decoding is done by applying \mathcal{M}^{-1} which recovers without error the typical set.

A Measure of Randomness

An ideal fair coin ($p_t = p_h$) is a source of perfect randomness, where all the events are equiprobable. We say that the source generates 1 *bit of randomness*, that will be denoted c in the following. Interestingly any biased source can be converted in a fair source (equiprobable). One way of doing it is by applying data compression to the output string. The $2^{nH(p)}$ typical sequences being equiprobable, assigning a string of $nH(p)$ bits to each typical sequence is a way of generating $nH(p)$ random bits. This gives an operational interpretation of the entropy as the rate of randomness generated by a large family of sources, as we will see in the next section.

5.2 Definitions

Two Families of Resources

In the following sections we are going to give an operational interpretation to different entropic quantities by showing how two partners can distill bits of correlations (denoted $[cc]$ ¹) from a noisy correlation (denoted $\{cc\}$). Subsequently we will study its relation with a protocol of communication of noiseless bits (denoted $[c \rightarrow c]$) through a noisy channel (denoted $\{c \rightarrow c\}$). Such protocols consider two different scenarios, generating either static or dynamic resources.

Static Resources

A static resource $\{cc\}$ consist in a noisy distribution $p(x, y)$ shared by Alice and Bob, as shown in Fig. 5.3. The task will be to convert the noisy distribution, through distillation, into bits of correlations $[cc]$.

¹In the following we will use the following notation for information (randomness) processing resources: a noiseless binary channel will be denoted by $\{c \rightarrow c\}$, and $[cc]$ for a noiseless bit of distributed correlation, reflecting their static/dynamical nature. A noisy correlated distribution is denoted $\{cc\}$, and a general noisy channel is denoted by $\{c \rightarrow c\}$.

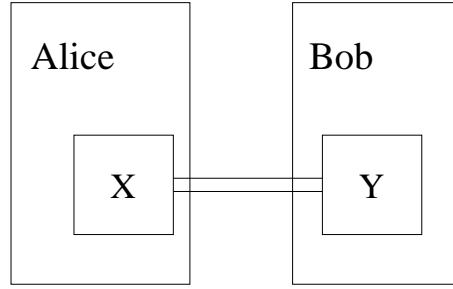


Figure 5.3: A static bipartite source consist in pairs of correlated letters (X, Y) according to the distribution $p(x, y)$ and shared by Alice (X) and Bob (Y).

Dynamic Resources

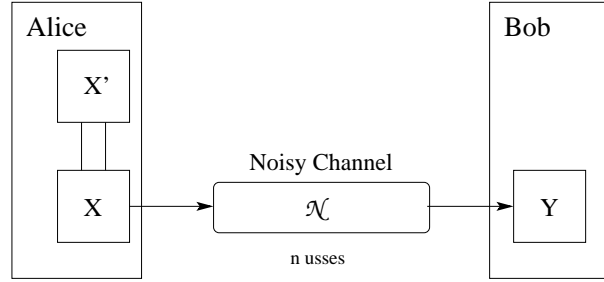


Figure 5.4: Alice holds a pure entangled state $|\psi\rangle_{AB_0}$. She sends to Bob the mode B_0 through the quantum channel \mathcal{N} .

A dynamic resource $\{c \rightarrow c\}$ is a noisy channel \mathcal{N} taking a bipartite perfectly correlated distribution $p(x, x') = p_x \delta_{x, x'}$ at Alice site, into a noisy distribution $p(x, y)$ shared between Alice and Bob, as shown in Fig. 5.4. The objective will be to transmit noiseless bits of correlations ($[c \rightarrow c]$) through a noisy channel ($\{c \rightarrow c\}$).

Information Theory

Shannon developed his theory in order to apply it to the compression and transmission of data through communication channels. One can see any text (or image, song...) as a generalized source:

Definition: A *source* consists in a sequence of random variables X_1, X_2, \dots whose values represent the output of the source taking values from the finite alphabet $\mathcal{H} = \{0, 1, \dots, d\}$, of size d .

Independent and Identically Distributed Sources

The coin is a simple example of a large family of sources, the so-called "independent and identically distributed" (i.i.d.) sources.

Definition: An *independent and identically distributed (i.i.d.) source* consists in a sequence of random variables X_1, X_2, \dots, X_n whose values represent the output of

the source taking values from the finite alphabet $\mathcal{H} = \{0, 1, \dots, d\}$, of size d . All variables X_i are independent $p(x_1, x_2, \dots) = p(x_1)p(x_2)\dots p(x_n)$ and identically distributed $X_i = X$ with probability mass function $p(x) = \Pr\{X = x\}, x \in \mathcal{H}$.

As in this thesis we are concerned with key distribution, which are nothing else than random and private data, restricting our study to i.i.d. sources will be sufficiently general.

Shannon work assumed that the text was resulting from an i.i.d. source, then compressing (transmitting) information or randomness are equivalent tasks. Real sources do not generally behave as i.i.d. sources, as it is easy to check that the letters in this English text do not occur in an independent fashion; strong correlations exist between them. Nevertheless, the assumption of an i.i.d. works already pretty well in practice, and the ideas introduced to deal with the special case of an i.i.d. source can be generalized to more sophisticated sources.

Shannon Entropy

The first example concerned a random source with an alphabet composed of two letters $\mathcal{A} = \{\text{"head"}, \text{"tail"}\}$. This can be generalized to random sources X with a larger alphabet $\mathcal{H} = \{0, 1, \dots, d\}$ where each letter x occurs with probability $p(x)$. In a string of n letters, x typically occurs about $np(x)$ times, the number of typical string being,

$$\frac{n!}{\prod (np(x))!} \approx 2^{nH(X)} \quad (5.7)$$

and all the typical sequences have the same probability $p(x) \approx 2^{-nH(X)}$, where $H(X)$ is the Shannon entropy

$$H(X) = - \sum_{x \in \mathcal{H}} p(x) \log p(x), \quad (5.8)$$

The Shannon entropy $H(X)$ gives the optimal data compression rate that can be reached for a given i.i.d. source, giving a measure of the randomness (or information) generated by the source. Alternatively, it can be seen as the amount of uncertainty about X before we learn its value.

Joint Entropy

The *joint entropy* $H(X, Y)$ of a pair of discrete random variables (X, Y) with alphabets $\mathcal{H} = \{0, 1, \dots, d_1\}$ and $\mathcal{J} = \{0, 1, \dots, d_2\}$ with a joint distribution $p(x, y)$ is defined as

$$H(X, Y) = - \sum_{x \in \mathcal{H}} \sum_{y \in \mathcal{J}} p(x, y) \log p(x, y). \quad (5.9)$$

As for the usual entropy $H(X)$ the joint entropy $H(X, Y)$ has an interpretation as the amount of uncertainty about the pair (X, Y) before we learn its value or as the optimal joint compression rate.

Conditional Entropy

The *conditional entropy* $H(Y|X)$ of a pair of discrete random variables (X, Y) with a joint distribution $p(x, y)$ is defined as

$$H(Y|X) = \sum_{x \in \mathcal{H}} p(x) H(Y|X = x), \quad (5.10)$$

$$= - \sum_{x \in \mathcal{H}} \sum_{y \in \mathcal{J}} p(x, y) \log p(y|x). \quad (5.11)$$

The conditional entropy $H(Y|X)$ has an interpretation as the average uncertainty we have about Y when we know X . Both joint and conditional entropies are related by the following *chain rule*, which can be recovered using a Venn diagram (Fig. 5.5).

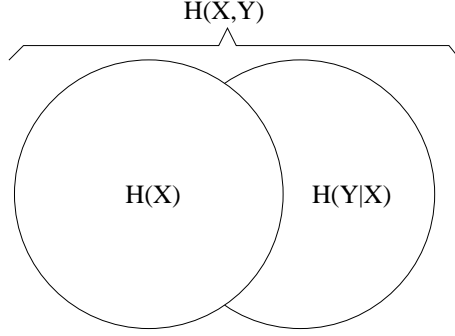


Figure 5.5: Relation between $H(X)$ and $H(Y|X)$.

Theorem 1 (Chain rule)

$$H(X, Y) = H(X) + H(Y|X). \quad (5.12)$$

Proof

Using the definition of the conditional probability one can write

$$\log p(x, y) = \log p(x) + \log p(y|x), \quad (5.13)$$

and take the expectation value of both sides of the equation to obtain the theorem. \square

The chain rule can be generalized to a collection of random sources X_1, X_2, \dots, X_n

$$H(X_1, X_2, \dots, X_n) = \sum_{i=1}^n H(X_i | X_{i-1}, \dots, X_1). \quad (5.14)$$

Relative Entropy

The *relative entropy* or *Kullback-Leibler distance* between two probability mass functions $p(x)$ and $q(x)$ is defined as

$$D(p \| q) = \sum_{x \in \mathcal{H}} p(x) \log \frac{p(x)}{q(x)}, \quad (5.15)$$

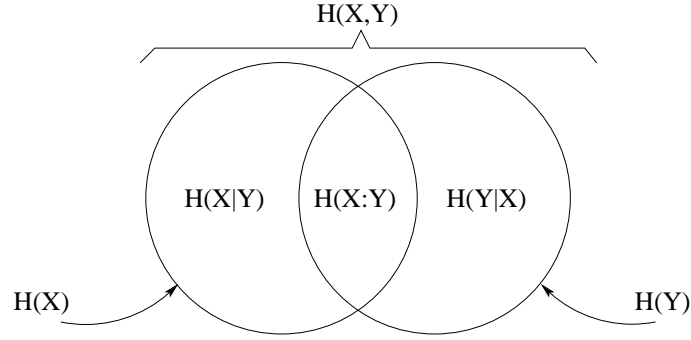
which is always non-negative and zero if and only if $p(x) = q(x)$.

Mutual Entropy

The *mutual information* $H(X:Y)$ of a pair of discrete random variables (X, Y) with a joint distribution $p(x, y)$ is defined as

$$H(X:Y) = - \sum_{(x,y) \in \mathcal{H} \times \mathcal{Y}} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} = D(p(x, y) \| p(x)p(y)). \quad (5.16)$$

The mutual entropy (or mutual information) $H(X:Y)$ of a bipartite source (X, Y) has an interpretation as the amount of correlations between Alice and Bob, measured in bits of correlation. Using the definition of the mutual information, conditional entropy and entropy one can easily derive the following relations:

Figure 5.6: Relation between $H(X)$, $H(Y|X)$, $H(X|Y)$ and $H(X,Y)$.**Theorem 2 (Mutual information and entropy)**

$$H(X:Y) = H(X) - H(X|Y) \quad (5.17)$$

$$H(X:Y) = H(Y) - H(Y|X) \quad (5.18)$$

$$H(X:Y) = H(X) + H(Y) - H(X,Y) \quad (5.19)$$

The relation between $H(X)$, $H(Y|X)$, $H(X|Y)$ and $H(X:Y)$ can be recovered using a Venn diagram (Fig. 5.6).

Conditional Mutual Entropy

The conditional mutual entropy (information) of random variables X and Y given Z is defined as

$$H(X:Y|Z) = H(X|Z) - H(X|Y, Z), \quad (5.20)$$

which satisfies the chain rule,

$$H(X_1, X_2, \dots, X_n:Y|Z) = \sum_{i=1}^n H(X_i:Y|X_{i-1}, \dots, X_1, Z). \quad (5.21)$$

If Z is completely decorrelated from the rest, it becomes the chain rule for the mutual information.

Properties of the Entropies

We now give some simple relations between the different entropies.

1. $D(p||q) \geq 0$
2. $H(X) \geq 0$.
3. $H(Y|X) \geq 0$.
4. $H(X) \leq \log d$.
5. Subadditivity of entropy: $H(X, Y) \leq H(X) + H(Y)$.
6. Concavity of $H(X)$: if $X = \sum p_i X_i \Rightarrow H(X) \geq \sum p_i H(X_i)$.
7. $H(X, Y) \geq \max\{H(X), H(Y)\}$.
8. $H(X:Y) \geq 0$.

9. $H(X:Y) \leq \min\{H(X), H(Y)\}$.
10. $H(X:Y|Z) \geq 0$.
11. Conditioning reduces entropy: $H(X|Y, Z) \leq H(X|Y)$.
12. For a fixed transition $p(y|x)$ $H(X:Y)$ is a concave function of $p(x)$.
13. For a fixed input $p(x)$ $H(X:Y)$ is convex in $p(y|x)$.

Proof

- (1) Using the concavity of the logarithmic function and Jensen's inequality, see [51] for a detailed proof.
- (2) $0 \leq p(x) \leq 1$ implies $-\log p(x) \geq 0$, with equality when the distribution is deterministic $p(x) = \delta_{x,i}$.
- (3) The proof is direct as by definition the conditional entropy is an average of entropies.
- (4) We define $u(x)$ as the uniform distribution over an alphabet \mathcal{X} of size d . The proof is easy using the definition $H(X) = \log |\mathcal{X}| - D(p(x) || u(x))$ and the non-negativity of the relative entropy.
- (5) $H(X) + H(Y) - H(X, Y) = H(X:Y) = D(p(x, y) || p(x)p(y)) \geq 0$.
- (6) Defining a joint source where Y encodes which source X_y is used gives $\sum p_y H(X_y) = H(X|Y)$. Using $H(X, Y) = H(Y) + H(X|Y) \leq H(X) + H(Y)$ gives the result.
- (7) Using (3) and $H(X, Y) = H(X) + H(Y|X) = H(Y) + H(X|Y)$.
- (8) equivalent to (5).
- (9) Using $H(X:Y) = H(X) - H(X|Y) = H(Y) - H(Y|X)$ and (3).
- (10) $D(p || q) \geq 0$ can be rewritten as $D(p(y|x) || q(y|x)) \geq 0$ which gives $H(X:Y|Z) = D(p(x, y|z) || p(x|z)p(y|z)) \geq 0$.
- (11) Using the chain rule we have $H(X|Y) = H(X|Y, Z) + H(X:Z|Y)$ and (10).
- (12) We define a third variable Z which encodes which transition $p_z(x)$ is used. Using $H(X, Z:Y) = H(X:Y) + H(Z:Y|X) = H(Z:Y) + H(X:Y|Z)$ and $H(Z:Y) \geq 0$ and $H(Z:Y|X) = H(Y|X) - H(Y|X, Z) = 0$ as $p(y|x)$ is independent of Z , gives the final result.
- (13) We define a third variable Z which encodes which input $p_z(y|x)$ is used. Using $H(X:Y, Z) = H(X:Z) + H(X:Y|Z) = H(X:Y) + H(X:Z|Y)$ and $H(X:Z|Y) \geq 0$ and $H(X:Z) = 0$ gives the result. \square

5.3 Entropy Operational Interpretation

The concept of typical sequence being at the core of all the results of information theory we rigorously generalize the concept of typical sequence beyond the binary case. Suppose X is an i.i.d. information source. Given $\epsilon > 0$ we say that a string of source symbols $x^n = x_1 x_2 \dots x_n$ is ϵ -typical ($\in T_X$) if

$$2^{-n(H(X)+\epsilon)} \leq p(x_1 x_2 \dots x_n) \leq 2^{-n(H(X)-\epsilon)} \quad (5.22)$$

A useful equivalent reformulation of the definition is

$$\left| \frac{1}{n} \log \frac{1}{p(x_1 x_2 \dots x_n)} - H(X) \right| \leq \epsilon. \quad (5.23)$$

Using the law of large numbers we can prove the following theorem [51].

Theorem of typical sequences

1. *Most of the sequences are typical:* Fix $\epsilon > 0$. Then for any $\delta > 0$, for sufficiently large n , the probability that a sequence is ϵ -typical is at least $1 - \delta$.
2. *Size of the typical set:* For any fixed $\epsilon > 0$ and $\delta > 0$, for sufficiently large n , the number $|T_X|$ of ϵ -typical sequences satisfies

$$(1 - \delta)2^{n(H(X) - \epsilon)} \leq |T_X| \leq 2^{n(H(X) + \epsilon)}. \quad (5.24)$$

3. *Small sets have zero probability:* Let $S(n)$ be a collection of size at most 2^{nR} , of length n sequences from the source, where $R < H(X)$ is fixed. Then for any $\delta > 0$ and for sufficiently large n ,

$$\sum_{x \in S(n)} p(x) \leq 2\delta. \quad (5.25)$$

The third point states that for large n the probability of a sequence output from the source lying in a subset $S(n)$ of size 2^{nR} goes to zero, as the number of typical sequences becomes exponentially larger than 2^{nR} .

Shannon noiseless channel coding theorem

A *compression scheme* of rate R maps possible sequences $x^n = (x_1, \dots, x_n)$ of n letters from a finite alphabet \mathcal{H} (containing d elements) to a bit string of length nR which we denote by $C^n(x^n) = C^n(x_1, \dots, x_n)$. The *decompression scheme* takes the nR compressed bits and maps them back to a string of n letters from \mathcal{H} . A compression-decompression scheme $D^n(C^n(x))$ (as in Fig. 5.7) is said to be *reliable* if $\Pr(D^n(C^n(x)) = x) \rightarrow 1$ as $n \rightarrow \infty$. The Shannon's noiseless coding theorem specifies for which rates R a reliable compression scheme exists.

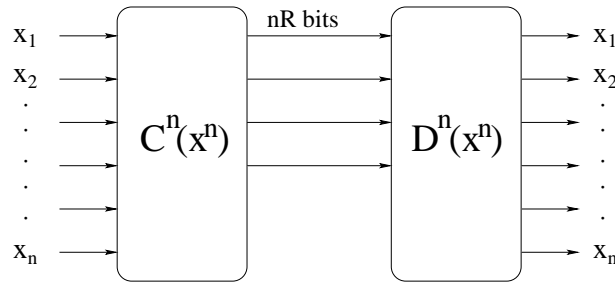


Figure 5.7: A compression scheme C^n of rate R maps possible sequences $x^n = (x_1, \dots, x_n)$ of letters from a finite alphabet \mathcal{H} to a bit string of length nR . The decompression scheme takes the nR compressed bits and maps them back to the string of the n original letters reliably.

Theorem 3 (Shannon's noiseless channel coding theorem) *Suppose X is an i.i.d. information source with entropy rate $H(X)$. Suppose $R > H(X)$. Then there exists a reliable compression scheme of rate R for the source. Conversely, if $R < H(X)$ then any compression scheme will not be reliable.*

Proof

The method of compression as described previously consists in checking if the output of the source is ϵ -typical. If it is not we output a failure sequence of nR zeros (this generates an error). If the output is typical then we compress the output simply by storing an index of the particular typical sequence using nR bits. Choosing $\epsilon > 0$ such that $H(X) + \epsilon < R$, for $\delta > 0$ and large n there are at most $2^{n(H(X)+\epsilon)} < 2^{nR}$ typical sequences. Then the probability of error reads,

$$P_e = \Pr(x \notin T_X) \leq \delta, \quad (5.26)$$

Conversely if $R < H(X)$ we know by the point (3) of the theorem of typical sequences that the probability that an output sequence lies in a subset $S(n)$ of size 2^{nR} goes to zero, for sufficiently large n . Thus any compression scheme with $R < H(X)$ cannot be reliable. \square

A measure of Randomness or Information

As we mentioned in the introduction, the entropy $H(X)$ is a measure of the randomness that can be generated by an i.i.d. source. The $2^{nH(p)}$ typical sequences being equiprobable, assigning a string of $nH(p)$ bits to each typical sequence is a way of generating $nH(p)$ random bits (c). In an i.i.d. model of source of information the entropy can be equivalently interpreted as the rate of information generated by the source.

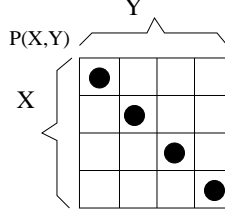
A measure of Correlations

Figure 5.8: A noiseless correlated distribution corresponds to a joint distribution $p(x, y)$, where once we know x (y) we have total certainty about y (x). The black dots correspond to pairs of the joint random variable (x, y) that have non-zero probability.

A bipartite source generates a noiseless correlated distribution when there is a bijection b that maps each element of X into an element of Y , as in Fig. 5.8. By knowing x^n (y^n) Alice (Bob) will know Bob's string y^n (Alice's string x^n) with full certainty. The joint distribution reads,

$$p(x, y) = p_x \delta_{y, b(x)}. \quad (5.27)$$

When the noiseless correlated source has equiprobable probability distribution $p(x, y) = \delta_{y, b(x)}/d$ we see that after distributing n pairs of symbols to Alice and Bob, both share $n \log d$ perfectly correlated bits, which locally look like $n \log d$ bits of randomness. We say that Alice and Bob share $n \log d$ bits of correlation (denoted $n \log d[cc]$).

Non-Uniform Noiseless Correlation For non-uniform noiseless distribution it is easy to check that Alice and Bob can extract $nH(X)$ bits of correlation from a string of n correlated pairs of letters drawn from the noiseless distribution $p(x, y)$, by applying a

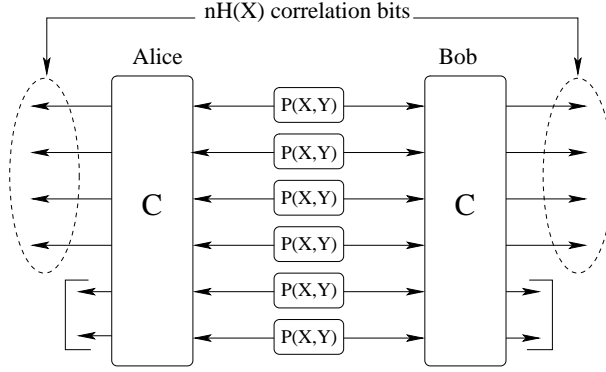


Figure 5.9: Alice and Bob can output $nH(X)$ shared random bits out of a noiseless correlated source (X, Y) by applying the compression map C without needing any communication.

data compression at each location (Alice and Bob), as shown in Fig. 5.9. This gives an operational interpretation of the entropy as the correlations generated by a noiseless correlated source.

Degenerate Noiseless Correlations

The source generates a correlated distribution where there is a mapping f from X to non-overlapping sets of Y (or vice-versa). The joint distribution then reads,

$$p(x, y) = p_x \delta_{x, f^{-1}(y)}. \quad (5.28)$$

By a local mapping $g(y)$ Bob can transform the degenerate noiseless correlated distribution between elements of X and non-overlapping subsets of Y to a noiseless distribution between element of X and elements of $g(Y)$, as shown in Fig. 5.10. Then considering that local operations cannot increase the correlations, the maximum bits of correlation that Alice and Bob can extract is $\min\{H(X), H(Y)\}$.

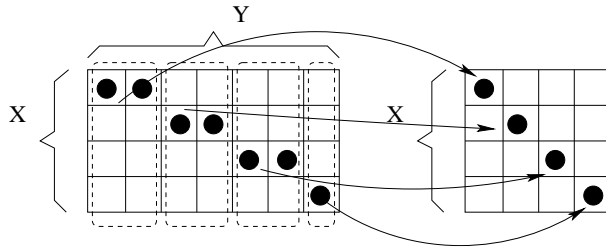


Figure 5.10: Applying a map $g(y)$ on Y Bob can transform the degenerate noiseless correlated distribution between elements of X and non-overlapping subsets of Y into a noiseless distribution.

Generalization This can be generalized to noisy correlated sources where the mutual entropy $I(X:Y)$ will be the measure of the amount of correlations that can be extracted from the joint distribution. Notice that for noiseless correlations $H(X:Y) = \min\{H(X), H(Y)\}$, which saturates the upperbound $H(X:Y) \leq \min\{H(X), H(Y)\}$ (property (9) of the entropies).

5.4 Data Merging and Correlation Distillation

In this section we are going to give an operational definition of the conditional entropy as the average amount of information that one partner has to send to another in order to allow him to get the full information about the bipartite sequence.

Jointly and Conditional Typical Sequences

Before giving an operational interpretation to the conditional entropy we need to introduce the joint and conditional typical sets.

Jointly Typical Set

The set $T_{X,Y}$ of jointly typical sequences $\{x^n, y^n\}$ with respect to the distribution $p(x, y)$ is the set of n -sequences with entropies ϵ -close to the true entropies, i.e.,

$$T_{X,Y} = \left\{ (x^n, y^n) \in \mathcal{X}^n \times \mathcal{Y}^n : p(x^n) = 2^{-(H(X) \pm \epsilon)}, \right. \quad (5.29)$$

$$\left. p(y^n) = 2^{-(H(Y) \pm \epsilon)}, p(x^n, y^n) = 2^{-n(H(X,Y) \pm \epsilon)} \right\}, \quad (5.30)$$

where

$$p(x^n, y^n) = \prod_{i=1}^n p(x_i, y_i), \quad (5.31)$$

and we used the notation $a^n = 2^{-n(b \pm \epsilon)}$ to mean

$$\left| \frac{1}{n} \log a^n - b \right| < \epsilon, \quad (5.32)$$

for n sufficiently large. In Fig. 5.11 we observe a net composed of the ensemble $T_X \otimes T_Y$ built by combining all typical sequences of X and Y . We observe that in general not all are jointly typical $T_{X,Y}$ (dark circles), as for example in perfectly correlated distributions (Fig. 5.8). The set $T_{X,Y}$ of jointly typical sequences satisfies

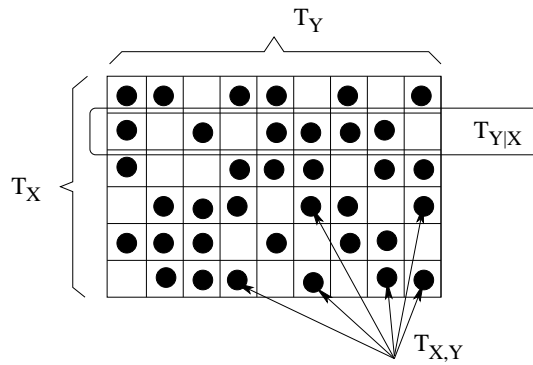


Figure 5.11: Each line (column) correspond to a typical sequence of T_X (T_Y). Each dark circle correspond to a typical sequence of $T_{X,Y}$. The dotted points in a given line (x^n) correspond to the set $T_{Y|x^n}$.

the following theorem, which is a generalization of the theorem of typical sequences.

Theorem 4 (Joint Asymptotic Equipartition Property) *For any $\epsilon > 0$, for sufficiently large n , and $S \subseteq \{X, Y\}$*

1. $\Pr(s^n \in T_S) \geq 1 - \epsilon.$
2. $s^n \in T_S \Rightarrow p(s^n) = 2^{-n(H(S) \pm \epsilon)}.$
3. $|T_S| = 2^{n(H(X) \pm 2\epsilon)}.$

Conditional Typical Set:

The set $T_{Y|x^n}$ of the set of sequences y^n that are jointly typical with a particular typical $x^n \in T_X$ is defined as,

$$T_{Y|x^n} = \left\{ (x^n, y^n) \in T_{X,Y} : p(y^n|x^n) = 2^{-n(H(Y|X=x^n) \pm \epsilon)} \right\}. \quad (5.33)$$

In Fig. 5.11 we observe that for each typical sequence of x^n there is a subset of typical sequences of Y that are at the same time jointly typical with X . This is just the ensemble $T_{Y|x^n}$.

Theorem 5 (Conditional Asymptotic Equipartition Property) *For any $\epsilon > 0$, and sufficiently large n , if $(x^n, y^n) \in T_{X,Y}$ then*

1. $\Pr(y^n \in T_{Y|X}) \geq 1 - \epsilon,$
2. $y^n \in T_{Y|X} \Rightarrow p(y^n|x^n) = 2^{-n(H(Y|X) \pm 2\epsilon)},$
3. $|T_{Y|X}| = 2^{n(H(Y|X) \pm 3\epsilon)},$

where we have used the notation $T_{Y|X}$ to express that we take the average over all the $T_{Y|x^n}$.

Data Merging

Given the joint sequence $x^n y^n$ generated by the source (X, Y) defined over alphabets of dimension d and distributed to Alice (X) and Bob (Y). Data merging is a protocol that allows Alice to send all the information that Bob lacks about the joint sequence $x^n y^n$. The most trivial solution is Alice sending directly the sequence x^n using $n \log d$ bits of communication. Alice can improve it by applying data compression to x^n (compression up to $nH(X)$ bits on average) and sending the compressed data to Bob through a noiseless channel. Fortunately we can do even better, as we prove in this section, where the optimal solution is shown to be $nH(X|Y)$.

Theorem 6 (Data Merging) *Suppose (X, Y) is an i.i.d. distributed source shared by Alice (X) and Bob (Y). Suppose $R > H(X|Y)$, then there exists a reliable protocol that transfers X from Alice to Bob using R bits of noiseless communication. Conversely, if $R < H(X|Y)$ then the transfer will not be reliable.*

In order to prove the lower bound we need the following result,

Theorem 7 (Data Transfer) *By sending one bit of communication we can increase the correlations at most by one bit.*

Intuitively we see that the best we can do is to send a bit through a noiseless channel in order to increase the correlations by one unit, this can be formally proven, see Appendix H.

The Lower Bound

Interestingly one can write any bipartite correlated distribution $p(x, y)$ as a noiseless correlated tripartite distribution by appending a third partner called the Reference (denoted W) that memorizes which sequence (x^n, y^n) is shared by Alice and Bob. The tripartite distribution reads,

$$p(x, y, w) = p(x, y)\delta_{(x, y), w}. \quad (5.34)$$

The proof is strikingly simple if we group Alice and the Reference into a single partner, see Fig. 5.12. Using the "data transfer" (Theorem 7) it is trivial to derive the following

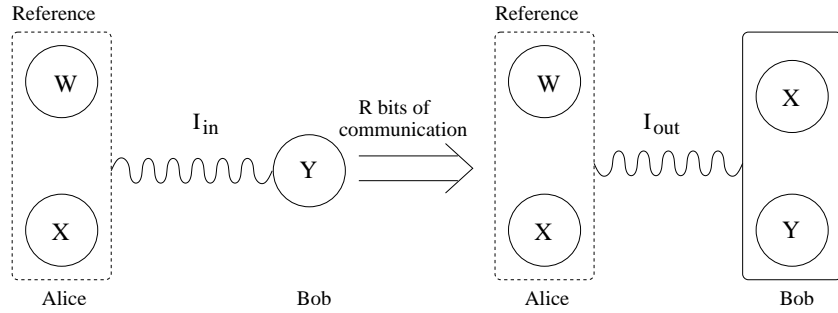


Figure 5.12: Alice sends R bits of communication to Bob in order to give Bob access to the joint distribution (X, Y) . Alice and the Reference are inside a dashed rectangle to emphasize that we have considered Alice and the Reference as a unique partner for technical reasons in the proof.

bound

$$I_{in} + R \geq I_{out}, \quad (5.35)$$

where I_{in} are the initial correlations between Alice+Reference and Bob and I_{out} are the final correlations after Alice has sent R bits to Bob through a noiseless channel. Using the fact that the tripartite distribution is perfectly correlated (5.34) and that knowing W then X is redundant, the initial amount of correlations reads

$$I_{in} = H(Y). \quad (5.36)$$

For a successful protocol, after Alice's communication of R bits, Bob has access to the distribution (X, Y) . It is then trivial to see that $I_{out} = H(X, Y)$, which together with equation (5.35) gives,

$$R \geq I_{out} - I_{in} = H(X, Y) - H(Y) = H(X|Y). \quad (5.37)$$

Achievability Proof

Now, we are going to present a protocol that succeeds to transfer X to Bob with a rate $R > H(X|Y)$, which shows that the lower bound derived previously is tight.

The Protocol

- *Encoding:* Alice independently generates a map $g(x^n)$ that assigns to each sequence $x^n \in T_X$ a bin (j) among a set of 2^{nR} bins, according to a uniform distribution.
- *Communication:* Alice sends to Bob the index of the bin (j) to which x^n belongs.

- *Decoding:* Given the received index j and his prior information y^n , Bob makes a guess $\hat{x}^n = f(j, y^n)$ of Alice data, where the mapping function f is defined only on the jointly typical set.

The protocol would be reliable if the average probability of error satisfies $P_e \rightarrow 0$ for $n \rightarrow \infty$. There are two sources of error on the protocol:

1. E_0 : The mapping function f being defined only on the set of jointly typical sets, if the pair (x^n, y^n) is not jointly typical we have an error.
2. E_1 : When for a given $y^n \in T_Y$ there is another jointly typical sequence on the same bin (j) as $\hat{x}^n = f(j, y^n)$.

The Joint Equipartition Theorem (Theorem 4) shows us that $\Pr(E_0) \rightarrow 0$ for sufficiently large n . Now we are going to prove that for sufficiently large n the bound $R \geq H(X|Y)$ can be asymptotically achieved, while satisfying $\Pr(E_1) \rightarrow 0$.

The probability $\Pr(E_1)$ is the probability that during the construction of the code using the map f Alice had assigned to the same bin two typical sequences of T_X which are jointly typical with the same sequence y^n ,

$$\begin{aligned} \Pr(E_1) &= \sum_{T_X \times T_Y} p(x^n, y^n) \Pr\{\exists \bar{x}^n \neq x^n : g(\bar{x}) = g(x), (x^n, y^n) \in T_{X,Y}\} \\ &\leq \sum_{T_X \times T_Y} p(x^n, y^n) \sum_{(x^n, \bar{x}^n) \in \{T_X|_{y^n} \setminus x^n\}} \underbrace{\Pr(g(\bar{x}) = g(x))}_{=2^{-nR}} \end{aligned} \quad (5.38)$$

$$= \sum_{T_X \times T_Y} p(x^n, y^n) 2^{-nR} |T_X|_Y \quad (5.39)$$

$$\leq 2^{-nR} 2^{n(H(X|Y)+3\epsilon)}, \quad (5.40)$$

which goes to zero for all $R > H(X|Y)$, for a sufficiently large n . \square

Correlations Distillation

Because data merging succeeds to transfer X^n from Alice to Bob using $nH(X|Y)$ bits of noiseless communication it allows Alice and Bob to extract $nH(X)$ bits of correlation (denoted $nH(X)[cc]$) from the joint distribution, as at the end of the protocol both partners know X^n perfectly. This can be summarized in the following resource inequality,

$$H(X|Y)[c \rightarrow c] + \{cc\} \geq H(X)[cc], \quad (5.41)$$

which express that by sending $nH(X|Y)$ bit of noiseless communication from Alice to Bob, both partners can extract $nH(X)$ bits of randomness out of n pairs of shared letters (x^n, y^n) generated by an i.i.d. probability distribution $p(x, y)$.

Strong Subadditivity

Using data merging we can get an operationally intuitive proof of strong subadditivity, which can be shown to be equivalent to *conditioning decrease entropy*

$$H(X|YZ) \leq H(X|Y). \quad (5.42)$$

Strong subadditivity is just the observation that if Bob has access to an additional register Z , then Alice surely does not need to send more information than in the case where Bob does not have access to Z . After all, Bob could always ignore Z .

5.5 Channel Capacity

The scheme studied in the previous section where Alice sends information to Bob in order to share a perfectly correlated string x^n (of size $nH(X)$ bits) can be transformed in a noiseless correlation distribution protocol (generating bits of correlation $[c \rightarrow c]$) through a noisy memoryless channel ($\{c \rightarrow c\}$), as we show below.

Correlation Distribution and Communication

The mathematical analog of a physical signaling (correlation distribution) system is shown in Fig. 5.13. A message W , drawn from the index set $\{1, 2, \dots, M\}$, is encoded by Alice in the sequence $x^n = f(W)$ using the encoding function $f : \{1, 2, \dots, M\} \rightarrow \mathcal{H}^n$. Alice sends the sequence x^n through the *memoryless channel* \mathcal{N} characterized by the probability transition function $p(y|x)$. Finally Bob guesses the index W by an appropriate decoding rule $\hat{W} = g(y^n)$. We define a (M, n) code as the ensemble of index $\{1, 2, \dots, M\}$ plus the encoding (f) and decoding functions (g). The rate of a code (M, n) is

$$R = \frac{\log M}{n} \text{ bits per transmission.} \quad (5.43)$$

A rate is said to be *achievable* if there exists a sequence of $(2^{nR}, n)$ codes such that the average error $P_e \rightarrow 0$ as $n \rightarrow \infty$.

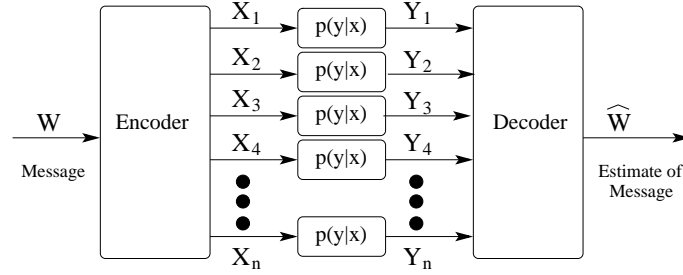


Figure 5.13: Alice encodes the message W in a sequence $x^n(W)$ and sends it through the memoryless channel \mathcal{N} ($p(y|x)$) which outputs a sequence y^n . Finally Bob guesses the index W by decoding y^n .

Memoryless Channel

In what follows we will restrict our study to noisy memoryless channels which are the channels that for a given input string $x^n = x_1x_2\dots x_n$ yields an output string $y^n = y_1y_2\dots y_n$, where each element y_i depends only on the input x_i ,

$$p(y^n|x^n) = \prod_{i=1}^n p(y_i|x_i). \quad (5.44)$$

The channel is then completely characterized by the transition probability $p(y|x)$.

Typical Sets Notice that if the characteristic $p(y|x)$ of the channel is known we can calculate $p(y)$, $p(x, y)$ and $p(x|y) = p(x, y)/p(y)$ from $p(y|x)$ and $p(x)$, which allows us to calculate all the typical sequences: T_{XY} , T_X , T_Y , $T_{Y|X}$ and $T_{X|Y}$.

Channel Capacity

The highest rate in bits per channel use at which correlations can be distributed (or information can be sent) with arbitrarily low probability of error is called the *channel capacity*, which is defined as

$$C = \max_{p(x)} H(X:Y), \quad (5.45)$$

where the maximum is taken over all possible input distributions $p(x)$. At the end of this section we shall show how to reach the capacity and that equation (5.45) is indeed the highest achievable rate.

Achievable Protocol

Alice's objective is to send a message W drawn from an equiprobable source $\{1, 2, \dots, M\}$ to Bob without any error through a memoryless noisy channel. Based on the data merging protocol presented in the previous section Alice can send information to Bob helped by a noiseless channel, using the following technique sketched in Fig. 5.14.

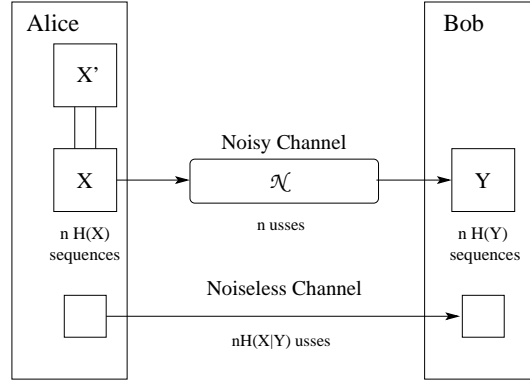


Figure 5.14:

Encoding Alice associates each message W to one typical sequences x^n generated by her source ($p(x)$). After making a copy x'^n of the generated sequence (that she keeps) Alice sends x^n to Bob through the noisy channel \mathcal{N} , Bob receiving a noisy version of it (y^n). Now we are back to the situation of the preceding section where Alice and Bob share a distributed source $p(x, y)$.

Error Correction Communication If Alice and Bob have access to a parallel noiseless channel, then Alice sending $nH(X|Y)$ bits to Bob will allow him to correct the errors in y^n , as shown in the preceding section, recovering x^n efficiently ($P_e \rightarrow 0$ when $n \rightarrow \infty$).

Rate Because at the end of the protocol Bob knows x^n without any error, Alice and Bob have succeed to share $nH(X)$ bits of correlations, or equivalently $nH(X)$ bits of communication. But in the process they have used $nH(X|Y)$ bits of communication through the noiseless channel, which is summarized in the following resource inequality (per signal sent),

$$H(X|Y)[c \rightarrow c] + \{c \rightarrow c\} \geq H(X)[c \rightarrow c], \quad (5.46)$$

which is the source-channel dual of the correlations distillation resource inequality (5.41). The net gain of correlations per signal sent through the noisy channel (denoted R) then reads,

$$R = H(X) - H(X|Y) = H(X:Y), \quad (5.47)$$

which is just the mutual entropy between Alice and Bob data.

Error Correcting Codes

In practice a noiseless channel does not exist, but interestingly one can get rid of the error correction communication through the noiseless channel. Alice instead of using the full ensemble of typical sequences T_X to encode W uses a smaller set of typical sequences of size $2^{nH(X:Y)}$ to encode W . This protocol is summarized by the following resource inequality (per signal sent),

$$\{c \rightarrow c\} \geq H(X:Y)[c \rightarrow c]. \quad (5.48)$$

The trick is to select among the $2^{nH(X)}$ typical sequences of T_X a code \mathcal{C} of size 2^{nR} so that the codewords are sufficiently far away one from the others such that Bob errors become negligible. Each typical sequence y^n received by Bob could have

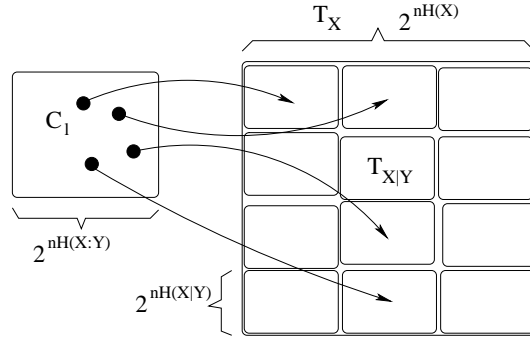


Figure 5.15: We construct a reliable (Bob estimation error being negligible) code \mathcal{C}_l of size $2^{nH(X:Y)}$ by dividing the typical space T_X into $2^{nH(X:Y)}$ non-overlapping cells ($T_{X|Y}$) of size $2^{nH(X|Y)}$ by selecting one single typical sequence from each cell.

come on average from one of the $2^{nH(X|Y)}$ sequences of the conditional typical subspace $T_{X|y^n}$ (remember that $H(X|Y)$ measure Bob's uncertainty on Alice data). By properly dividing the typical set T_X into 2^{nR} non-overlapping cells of size $2^{nH(X|Y)}$ and selecting one single codeword from each cell, we construct a reliable ($P_e \rightarrow 0$ when $n \rightarrow \infty$) code \mathcal{C} of rate R , as shown in Fig. 5.15. The size of the space T_X being $2^{nH(X)}$, we see that the optimal rate R we can achieve is

$$R = \frac{1}{n} \log \frac{|T_X|}{|T_{Y|X}|} = H(X) - H(X|Y) = H(X:Y). \quad (5.49)$$

This shows that Alice can distribute $H(X:Y)$ bits of correlation using a noisy channel. Its easy to see that T_X can be divided in $2^{H(X|Y)}$ different non-overlapping subspaces, each one corresponding to a given code \mathcal{C}_l , which allows us to interpret the $nH(X|Y)$ bits sent to Bob by Alice during the data merging protocol as telling Bob in "which code \mathcal{C}_l " the sequence x^n lies.

Optimality

Let us prove that the channel capacity (5.45) is indeed the optimal rate that can be achieved by a protocol which gives an arbitrarily low probability of error. Assume we have a code $(2^{nR}, n)$ with zero probability of error, the decoder output $\hat{W} = g(Y^n)$ (see Fig. 5.13) is equal to the input index W with probability 1, then $H(W|Y^n) = 0$.

We can now write

$$nR = nH(X^n(W)) = \underbrace{H(X^n|Y^n)}_{=0} + H(X^n : Y^n) \quad (5.50)$$

$$= H(Y^n) - H(Y^n|X^n) \quad (5.51)$$

$$\stackrel{(a)}{=} H(Y^n) - \sum_{i=1}^n H(Y_i|Y_1, \dots, Y_{i-1}, X^n) \quad (5.52)$$

$$\stackrel{(b)}{=} H(Y^n) - \sum_{i=1}^n H(Y_i|X_i) \quad (5.53)$$

$$\stackrel{(c)}{\leq} \sum_{i=1}^n H(Y_i) - \sum_{i=1}^n H(Y_i|X_i) \quad (5.54)$$

$$= \sum_{i=1}^n H(X_i : Y_i) \quad (5.55)$$

$$\stackrel{(d)}{\leq} nC, \quad (5.56)$$

where (a) follows from the chain rule of conditional entropy, (b) from the definition of a memoryless channel (Y_i depends only on X_i), (c) from the subadditivity of the entropy and (d) from the definition of C . Hence for any zero-error code, C gives the optimal communication rate for a fixed channel.

5.6 Decorrelation

In the previous section we have given an operational interpretation of the mutual information as the amount of communication (or correlation distribution) that can be achieved with a given channel and source. Here we show that mutual information is a measure of the correlations inside a bipartite distribution, giving the amount of randomness needed to decorrelate a bipartite distribution.

Consider the following scenario, Alice and Bob share a correlated bipartite source (X, Y) that generates the distribution $p(x, y)$ ($H(X:Y) \neq 0$). In addition Alice has access to a second source W , which is decorrelated from the bipartite source ($p(x, y, w) = p(x, y)p(w)$). In order to decorrelate her data (X) from Bob's data (Y) Alice combines the sequence w^n with x^n using a deterministic function $(x'^n) = f(x^n, w^n)$ that preserves Alice's probability distribution, $p(x') = p(x)$.

Proof To simplify the proof we append to each independent source (X, Y) and W a perfectly correlated reference system R_1 and R_2 , respectively, as shown in Fig 5.16. Both reference systems being independent at the beginning of the process, the initial amount of correlation per signal between the reference systems (R_1, R_2) and (X, Y, W) reads,

$$I_{in} = H(R_1) + H(R_2) = H(W) + H(X, Y). \quad (5.57)$$

After applying the function f to systems X and W and discarding W' , if Alice and Bob distributions are independent, so are their references R'_1 and R'_2 . The final amount of correlations reads,

$$I_{out} = H(R'_1) + H(R'_2) = H(X) + H(Y). \quad (5.58)$$

Because local operations (applying f) and discarding a system (W') can only decrease the correlations ($I_{in} \geq I_{out}$) we obtain a lower bound for the amount of randomness

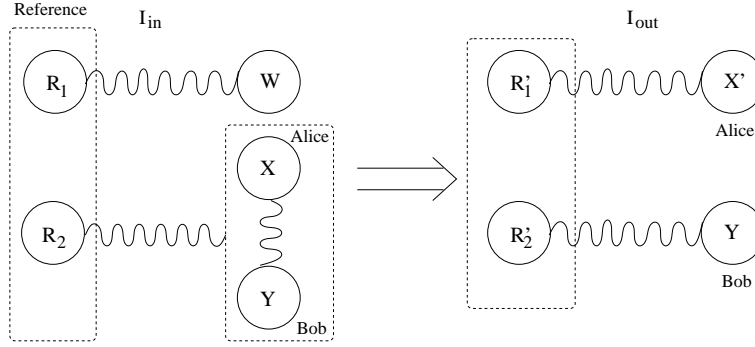


Figure 5.16: Alice and Bob share a correlated distribution (X, Y) which is perfectly correlated with the reference R_2 . The independent source of random bits W is perfectly correlated with R_1 . After Alice applying the function $X' = f(X, W)$ and discarding the system W the correlations between the references (R_1, R_2) and the rest of the systems can only decrease.

needed per signal $(H(W))$ for a successful decorrelation,

$$H(W) \geq H(X) + H(Y) - H(X, Y) = H(X:Y). \quad (5.59)$$

One can prove that this bound can be achieved applying the following protocol. Alice divides her typical set T_X into $2^{nH(X|Y)}$ codes C_l of size $2^{nH(X:Y)}$. At the beginning of the protocol Bob's uncertainty is on which code C_l lies x^n in. In order to increase Bob's uncertainty to the ensemble of the typical set T_X , Alice applies to each code C_l the following random permutation: Firstly, assign to each code a binary number c of size $nH(X:Y)$; Secondly, apply the XOR operation $c' = c \oplus r$, where r is a sequence of $nH(X:Y)$ random bits. Finally apply the mapping $c \rightarrow c'$, which applies a random permutation among the sequences of each code C_l .

5.7 Continuous Variables

Because in this dissertation we are interested in applications using continuous variables we need to introduce the concept of *differential entropy* $H(X)$, which is the entropy of a continuous random variable X with density $f(x)$,

$$H(X) = - \int_S f(x) \log f(x) dx, \quad (5.60)$$

where S is the support of the random variable. The size of the typical set $|T_X|$ now becomes a volume. Differential entropy is also related to the shortest description length, and shares most of the properties of the entropy of discrete random variables. One can similarly define conditional, mutual and relative differential entropies, which together with the differential entropy have exactly the same properties as those introduced in Section 5.2 for the discrete distributions.

There is just one but very important difference, the differential entropy is defined up to an arbitrary constant depending on the scaling. If we apply a rescaling $y = ax$ where $f_Y(y) = f_X(y/a)/|a|$, the differential entropy reads,

$$\begin{aligned} H(aX) &= - \int f_Y(y) \log f_Y(y) dy = - \int \frac{1}{|a|} f\left(\frac{y}{a}\right) \log \left(\frac{1}{|a|} f\left(\frac{y}{a}\right) \right) dy \\ &= - \int f_X(x) \log f_X(x) dx + \log |a| = H(X) + \log |a|. \end{aligned} \quad (5.61)$$

The conditional entropy being also ill defined. Interestingly, this illness in the definition vanishes when we consider a quantity defined as the difference of two entropies, such as for mutual information $H(X:Y) = H(Y) - H(Y|X)$, since as both arbitrary constants cancel out. Then quantities such as capacity or secret key distribution rate over a continuous variable channels are well defined.

Gaussian Distributions

We now introduce the important case of Gaussian distributions that we will use later in this dissertation.

Entropy Consider a normal distribution

$$g(x) = \frac{1}{\sqrt{2\pi V_X}} e^{-x^2/2V_X}. \quad (5.62)$$

with variance V_X . The differential entropy reads,

$$\begin{aligned} H(X) &= - \int g(x) \ln g(x) dx = - \int g(x) \left[-\frac{x^2}{2V_X} - \ln \sqrt{2\pi V_X} \right] dx \\ &= \frac{1}{2} + \frac{1}{2} \ln 2\pi V_X = \frac{1}{2} \log V_X + C, \end{aligned} \quad (5.63)$$

where C is an arbitrary constant changing with the scaling.

For a bipartite normal distribution with covariance matrix

$$K_{AB} = \begin{bmatrix} \langle x^2 \rangle & \langle xy \rangle \\ \langle xy \rangle & \langle y^2 \rangle \end{bmatrix} \quad (5.64)$$

the differential entropy reads [51],

$$H(X, Y) = \frac{1}{2} \log(\det K_{AB}) + C', \quad (5.65)$$

and similarly for more parties.

Conditional Entropy The conditional entropy $H(Y|X)$ of the distribution Y conditioned on X it can be written as,

$$H(Y|X) = \int dx H(Y|X = x). \quad (5.66)$$

Using the definition of conditional probability $f(y|x) = f(x, y)/f(x)$ [133] read,

$$H(Y|X) = \frac{1}{2} \log V_{Y|X} + C, \quad (5.67)$$

where $V_{Y|X}$ is the variance of Y when X is known,

$$V_{Y|X} = \frac{\det K_{AB}}{V_X} = \langle y^2 \rangle - \frac{\langle xy \rangle^2}{\langle x^2 \rangle}. \quad (5.68)$$

Mutual Entropy The mutual entropy (mutual information) of a bipartite distribution has three equivalent definitions,

$$H(X:Y) = H(Y) - H(Y|X) = \frac{1}{2} \log \left[\frac{V_Y}{V_{Y|X}} \right] \quad (5.69)$$

$$= H(X) - H(X|Y) = \frac{1}{2} \log \left[\frac{V_X}{V_{X|Y}} \right] \quad (5.70)$$

$$= H(X) + H(Y) - H(X, Y) = \frac{1}{2} \log \left[\frac{V_X V_Y}{\det K_{AB}} \right]. \quad (5.71)$$

Chapter 6

Quantum Information Theory

6.1 Introduction

Quantum information theory may be understood in terms of interconversion between various resources, such as entanglement, classical correlations or secret keys. In this chapter we present the distillation and distribution through quantum channels of the first two resources. The issue secret key will be studied in more detail in the next chapter. In the following we will use capital letters (A) to refer to quantum states, where as we will use small letters (a) for the output of a measurement over the corresponding quantum system (A).

Two Related Families of Resources

In the previous chapter we have shown how two partners could distill bits of correlations (a resource denoted $[cc]$) from a noisy bipartite probability distribution (a resource denoted $\{cc\}$)¹ and its relation with a protocol of communication of noiseless bits (resource denoted $[c \rightarrow c]$) through a noisy channel (resource denoted $\{c \rightarrow c\}$). By analogy with the classical information theory in its quantum counterpart we define a noiseless unit of entanglement or e-bit ($(|00\rangle + |11\rangle)/\sqrt{2}$) which is denoted $[qq]$ and noiseless qubit channel denoted $[q \rightarrow q]$. Similarly as for noiseless resource, there are two different types of resource, either static or dynamic.

Static Resources

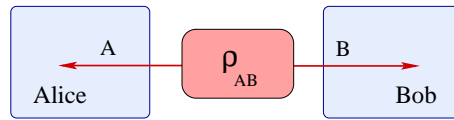


Figure 6.1: A static source distributes a quantum state ρ_{AB} between Alice and Bob, which can distill entanglement or classical correlation from it.

A static source $\{qq\}$ consist in a bipartite noisy quantum state ρ_{AB} shared by Alice and Bob, as shown in Fig. 6.1. The task will be to convert the quantum state, through distillation, into either static resources such e-bits (Bell pairs) $[qq]$, classical correlations $[cc]$ or secret bits $[ss]$.

¹Remember that $\{ \}$ represents a noisy resource and $[\]$ a noiseless one.

Dynamic Resources

A dynamic resource $\{q \rightarrow q\}$ consist in a noisy quantum channel \mathcal{N} taking a bipartite pure state $|\psi\rangle_{AB_0}$ at Alice site, into the mixed state $\rho_{AB} = (\mathbb{I}_A \otimes \mathcal{N})|\psi\rangle\langle\psi|_{AB_0}$ shared between Alice and Bob, as shown in Fig. 6.2. The quantum channel can be used to distribute either e-bits (or communicate qubits) $[q \rightarrow q]$, classical correlations $[c \rightarrow c]$ or secret bits $[s \rightarrow s]$.

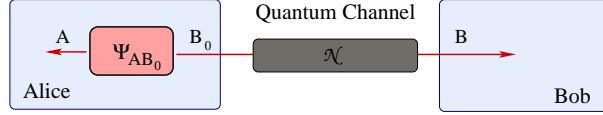


Figure 6.2: Alice holds a pure entangled state $|\psi\rangle_{AB_0}$ and sends mode B_0 to Bob through the quantum channel \mathcal{N} .

Independent and Identically Distributed Sources

The generalization of the independent and identically distributed (i.i.d.) classical source plays also a very important role in quantum information theory. In order to distribute e-bits ($[q \rightarrow q]$) between two partners, we need sources that generate entanglement.

Definition: An *i.i.d. entanglement source* generates independent entangled states, where the entire bipartite message state reads $|\psi\rangle_{AB} \otimes \dots \otimes |\psi\rangle_{AB}$, where

$$|\psi\rangle_{AB} = \sum_x \sqrt{p(x)} |x\rangle_A |x\rangle_B. \quad (6.1)$$

If the resources we are interested to generate are classical correlations ($[cc], [c \rightarrow c]$) or secret bits, we will then use the so called *classical-quantum sources* (C-Q sources).

Definition: An *i.i.d. Classical-quantum source* generates independent pairs of classical-quantum states, where the entire joint state of the classical register \mathbf{a} and the quantum signal \mathbf{B} has the density matrix $\rho_{aB} \otimes \dots \otimes \rho_{aB}$. For a source generating pure states ρ_{aB} we have,

$$\rho_{aB} = \sum_a p(a) |a\rangle\langle a| \otimes |\varphi_a\rangle\langle\varphi_a|_B. \quad (6.2)$$

Relation between both sources

Entanglement and classical-quantum sources are indeed related. A classical-quantum source ρ_{aB} such as (6.2) can be generated starting from an entanglement source $|\psi\rangle_{AB} = \sum_a \sqrt{p(a)} |a\rangle |\varphi_a\rangle$. Alice applies a projective measurement $\sum_a |a\rangle\langle a|$ on system A , obtaining the outcome a distributed according to the probability distribution $p(a)$ which projects Bob's state to $|\varphi_a\rangle_B$, as shown in Fig. 6.3.

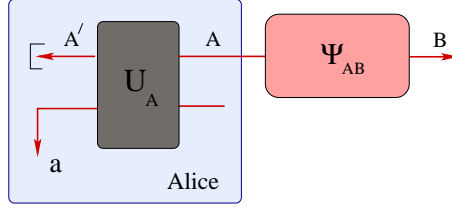


Figure 6.3: Alice's measurement of system A of the bipartite state ρ_{AB} , giving the result a and projecting B to ρ_B^a . Equivalently, a denotes the internal state of a preparer who prepares the state ρ_B^a according to a .

6.2 Entropies

Von Neumann Entropy

Von Neumann entropy of a quantum state defined in a Hilbert space \mathcal{H}_d of dimension d is defined by the formula,

$$S(\rho) = -\text{Tr}(\rho \log \rho), \quad (6.3)$$

where the logarithms are taken in base 2 as usual. If we choose the orthogonal basis $(|i\rangle)$ that diagonalizes ρ

$$\rho = \sum_i \lambda_i |i\rangle\langle i|, \quad (6.4)$$

then the von Neumann entropy can be re-expressed as the Shannon entropy of the eigenvalues (λ_i) of ρ ; $S(\rho) = H(\lambda)$. The von Neumann entropy can be interpreted as the smallest Hilbert space $\mathcal{H}_{S(\rho)}$ to which the quantum state ρ can be compressed reliably, as we show later in this chapter.

Extremum It is easy to show that the von Neumann entropy is minimal ($S(\rho) = 0$) when the state is pure $\rho = |\psi\rangle\langle\psi|$ and it is maximum ($S(\rho) = \log d$) when the state is maximally mixed $\rho = \mathbb{I}/d$.

Block-diagonal density matrix: States for which the density matrix is block diagonal are characterized by a probability distribution p_i and state ρ_i that have support on orthogonal subspaces

$$\rho = \sum_i p_i \rho_i. \quad (6.5)$$

It is then easy to prove that the von Neumann entropy of ρ reads,

$$S(\rho) = H(p) + \sum_i p_i S(\rho_i). \quad (6.6)$$

Proof

Let λ_i^j and $|e_i^j\rangle$ be the eigenvalues and corresponding eigenvectors of ρ_i . Observe that $p_i \lambda_i^j$ and $|e_i^j\rangle$ are eigenvalues and eigenvectors of $\sum_i p_i \rho_i$, and thus

$$S\left(\sum_i p_i \rho_i\right) = -\sum_{ij} p_i \lambda_i^j \log p_i \lambda_i^j \quad (6.7)$$

$$\stackrel{(a)}{=} -\sum_i p_i \log p_i - \sum_i p_i \sum_j \lambda_i^j \log \lambda_i^j \quad (6.8)$$

$$= H(p) + \sum_i p_i S(\rho_i). \quad (6.9)$$

where (a) follows from $\text{Tr}[\rho_i] = \sum_j \lambda_i^j = 1$.

Diagonal density matrix : For a probabilistic mixture of orthogonal states $|i\rangle$,

$$\rho = \sum_i p_i |i\rangle\langle i| \quad (6.10)$$

the von Neumann entropy is just the Shannon entropy of the diagonal terms $S(\rho) = H(p)$.

Remarks For sake of simplification, we will use the term *entropy* for the von Neumann entropy. In what follows we will use two different notations either $S(\rho_A)$ to stress that the entropy is a function of a given density matrix or $S(A) = S(\rho_A)$ to stress the role played by the system A .

Quantum Joint Entropy

The *joint entropy* $S(A, B)$ of a bipartite quantum system (A, B) with density matrix ρ_{AB} is defined as

$$S(A, B) = -\text{Tr}(\rho_{AB} \log \rho_{AB}). \quad (6.11)$$

As for the usual entropy $S(A)$ the joint entropy $S(A, B)$ has an interpretation as the optimal joint compression rate we can achieve on system AB .

Quantum Conditional Entropy

The *conditional entropy* $S(A|B)$ of a bipartite quantum system (A, B) with density matrix ρ_{AB} cannot be generalized directly from its classical counterpart using log functions, as there is no quantum generalization of the conditional probability. Nevertheless, one can define the quantum conditional entropy using the chain rule,

$$S(A, B) = S(B) + S(A|B), \quad (6.12)$$

giving

$$S(A|B) = S(A, B) - S(B). \quad (6.13)$$

Contrary to its classical counterpart the quantum conditional entropy can be negative.

Example This can be trivially shown for Bell states, such as $|\Phi^+\rangle_{AB} = [|00\rangle + |11\rangle]/\sqrt{2}$, where $S(A|B) = -1$, which can be proven from: (i) $S(A, B) = 0$ since the joint state is pure; (ii) $S(A) = 1$ since $\rho_A = \text{Tr}_B[|\Phi^+\rangle_{AB}] = \mathbb{I}/2$.

Interpretation The negativity of the conditional entropy is intimately related to the existence of entanglement, and it will play an important role in studying the distillation and distribution of entanglement.

Classical-Quantum Conditional Entropy

If the bipartite state can be written as a block diagonal density matrix

$$\rho_{aB} = \sum_a p(a) |a\rangle\langle a| \otimes \rho_B^a, \quad (6.14)$$

then the conditional entropy $S(B|a)$ reads,

$$S(B|a) = \sum_a p(a) S(\rho_B^a). \quad (6.15)$$

In this case the conditional entropy being an average of von Neumann entropies it is clearly non-negative. The C-Q conditional entropy plays an important role in the distillation and distribution of classical correlations over quantum channels.

Quantum Relative Entropy

As on Shannon information theory it is extremely useful to define a quantum version of the relative entropy. Suppose ρ and σ are density operators. The *relative entropy* is defined by

$$D(\rho || \sigma) = \text{Tr}[\rho \log \rho] - \text{Tr}[\rho \log \sigma], \quad (6.16)$$

which is always non-negative (Klein's Inequality [119, 136]) and zero if and only if $\rho = \sigma$.

Mutual Entropy

The *mutual entropy* (or *mutual information*) $S(A : B)$ of a bipartite quantum system (A, B) with density matrix ρ_{AB} must also be defined through an equation using only von Neumann entropies,

$$S(A:B) = S(A) + S(B) - S(A, B), \quad (6.17)$$

or using the relative entropy,

$$S(A:B) = D(\rho_{AB} || \rho_A \otimes \rho_B). \quad (6.18)$$

where $\rho_A = \text{Tr}_B \rho_{AB}$ and $\rho_B = \text{Tr}_A \rho_{AB}$.

Example 1: Decorrelated State A bipartite decorrelated state $\rho \otimes \sigma$ has null mutual information $S(A:B) = 0$. This can be trivially shown using the relative entropy definition of the mutual entropy and its property " $D(\rho || \sigma)$ is zero if and only if $\rho = \sigma$ ".

Example 2: Perfect Classical Correlations The quantum state $\rho_{AB} = \sum_i \frac{1}{d} |i\rangle\langle i|_A \otimes |i\rangle\langle i|_B$ being equivalent to a classical distribution with perfect classical correlations gives $S(A:B) = \log d$.

Example 3: Maximally Entangled State For a maximally entangled state $|\Psi\rangle_{AB} = \sum_i \sqrt{\frac{1}{d}} |i\rangle_A \otimes |i\rangle_B$ we have $S(A:B) = 2 \log d$ as $S(A) = S(B) = \log d$ ($\rho_A = \rho_B$ is maximally mixed) and $S(A, B) = 0$ ($|\Psi\rangle_{AB}$ being pure).

Interpretation Similarly as for probability distributions, the mutual entropy $S(A:B)$ of a bipartite source (A, B) has an interpretation as the minimal amount of randomness that is needed in order to decorrelate the quantum state ρ_{AB} . Using $S(A:B)$ random bits one can transform ρ_{AB} into $\rho'_{AB} = \rho_A \otimes \rho_B$, as shown in [93]. For example a Bell pair can be decorrelated using 2 random bits by applying a random bit flip followed by a random phase flip on one side. There is a full quantum scenario where the $S(A:B)$ bits of randomness are replaced by $S(A:B)/2$ e-bits (Bell pairs), as shown in [1]. For example, in order to decorrelate a Bell pair $(|00\rangle + |11\rangle)/\sqrt{2}$ a second pair is needed ($S(A:B)/2 = 1$). In order to implement the decorrelation Alice applies a noisy version of entanglement swapping without communicating the measurement result.

Conditional Mutual Entropy

The quantum conditional mutual entropy (information) of a bipartite quantum systems A and B given C is defined as

$$S(A:B|C) = S(A|C) - S(A|B, C), \quad (6.19)$$

which is non-negative because of the strong subadditivity of the entropy. If C is completely decorrelated from the rest, it becomes the usual mutual information $S(A:B)$.

Properties of the Entropies

We now give some simple relations between the different entropies.

1. $S(A) \geq 0$, with equality if the state is pure.
2. $S(A) \leq \log d$, with equality for a maximally mixed state.
3. $S(A|B) \geq -S(B)$.
4. Subadditivity of entropy: $S(A, B) \leq S(A) + S(B)$.
5. Concavity of $S(\rho)$: $S(\sum p_i \rho_A^i) \geq \sum p_i S(\rho_A^i)$.
6. $S(A) = S(B)$ if $S(A, B) = 0$.
7. $S(A, B) \geq |S(B) - S(A)|$.
8. $S(A:B) \geq 0$.
9. $S(A:B) \leq S(A) + S(B)$.
10. Conditioning reduces entropy: $S(A|B, C) \leq S(A|B)$.
11. $S(A:B|C) \geq 0$.
12. Discarding a quantum system never increases the mutual information:
 $S(A:B) \leq S(A:B, C)$.
13. $S(A, B|C) \leq S(A|C) + S(B|C)$.
14. Subadditivity of conditional entropy:
 $S(A_1, A_2|B_1, B_2) \leq S(A_1|B_1) + S(A_2|B_2)$.

Proof

(1) It follows from the definition of von Neumann entropy that is equal to the Shannon entropy of the eigenvalues of ρ_A . $H(\lambda) = 0$ if and only if just one eigenvalue is strictly positive which implies that the state must be pure.

(2) Let ρ_A be defined in a Hilbert space \mathcal{H}_d of dimension d . The proof is easy using the property $S(A) = \log |\mathcal{H}_d| - D(\rho_A || \mathbb{I}/d)$ and $D(\rho || \sigma) \geq 0$. We have equality if and only if $D(\rho_A || \mathbb{I}/d) = 0$ which occurs only if $\rho_A = \mathbb{I}/d$.

(3) Follows from the definition of the conditional entropy and (1): $S(A, B) \geq 0$.

(4) Similarly as in the classical case: $S(A) + S(B) - S(A, B) = S(A:B) = D(\rho_{AB} || \rho_A \otimes \rho_B) \geq 0$.

(5) Defining a joint state $\sum_i p(b) |b\rangle\langle b| \otimes \rho_A^b$ where b encodes which state ρ_A^b is used gives $\sum p_b S(\rho_A^b) = H(A|b)$. Using $H(A, B) \leq H(A) + H(B)$ gives the result.

(6) If $S(A, B) = 0$ the bipartite state is pure $\rho_{AB} = |\psi\rangle\langle\psi|_{AB}$. Then using the Schmidt decomposition $|\psi\rangle_{AB} = \sum_i \sqrt{p(i)} |i\rangle_A |i\rangle_B$, we see that the partial traces ρ_A and ρ_B have the same eigenvalues ($p(i)$) which implies the result.

(7) Introduce a system R that purifies A, B . Then (6) implies $S(R) = S(AB)$ and $S(RA) = S(B)$, which combined with (4) $S(R) + S(A) \geq S(RA)$, gives $S(A, B) \geq S(B) - S(A)$. Replacing A by B in the previous development gives the symmetric counterpart. Combining both gives the result.

(8) It is equivalent to (4).

(9) Using $S(A:B) = S(A) + S(B) - S(A, B)$ and $S(A, B) \geq 0$.

(10) This will be proved later. Intuitively it relies on the observation that if Bob has access to an additional register C , then Alice surely does not need to send more information than in the case where Bob does not have access to C .

(11) Using $S(A:BC) = S(A|C) - S(A|BC)$ and (10).

(12) $S(A:BC) = S(A:B) + S(A:C|B)$ and (11).

(13) $S(A, B|C) = S(A|C) + S(B|C) - (A:B|C)$ and (11).

(14) Use (13) followed by (10). \square

6.3 Quantum Data Compression

Consider an i.i.d. C-Q source generating a message of n pure quantum states $(\{p(x), |x\rangle\})$ with the density matrix of the entire message reading $\rho^n = \rho \otimes \dots \otimes \rho$. A compression of rate R for this source consists in a compression operation \mathcal{C}^n taking states from $\mathcal{H}^{\otimes n}$ to states in a 2^{nR} -dimensional Hilbert space \mathcal{H}_c^n , composed of nR qubits. The decompression operation \mathcal{D}^n restores the initial message from the compressed state. We say that the operation $\mathcal{D}^n \circ \mathcal{C}^n$ is reliable if the fidelity of $\mathcal{D}^n \circ \mathcal{C}^n$ approaches 1 in the limit of large n .

Entanglement-based Description

As explained in Appendix A, any quantum system Q prepared in a quantum state ρ with eigenvalues λ_x can be seen as the partial trace of a pure state

$$|\psi\rangle_{PQ} = \sum_x \lambda_x |x\rangle_P |x\rangle_Q, \quad (6.20)$$

where P is a reference system. Any quantum operation \mathcal{E} on system Q can be re-defined as the operation $I_P \otimes \mathcal{E}_Q$ applied to the state $|\psi\rangle_{PQ}$. How well the entanglement between Q and P is preserved is quantified by the *entanglement fidelity* $F(\rho, \mathcal{E})$ which is a function of ρ and \mathcal{E} defined by

$$F(\rho, \mathcal{E}) = \langle \psi | [(I_P \otimes \mathcal{E}_Q) |\psi\rangle \langle \psi|] | \psi \rangle. \quad (6.21)$$

One can then present the quantum compression \mathcal{C}^n as the operation that compresses the state $\rho^{\otimes n}$ of system Q defined over the space $\mathcal{H}^{\otimes n}$ to a 2^{nR} -dimensional typical subspace Q' , see Fig. 6.4, while preserving its entanglement with system P .

In order to prove the quantum version of the noiseless coding theorem we need to generalize the concept of typical sequences to the quantum scenario. We are going to show that for large n , the density matrix $\rho^{\otimes n}$ has nearly all its support on a subspace of $\mathcal{H}^{\otimes n}$ called the typical subspace of dimension $2^{nS(\rho)}$.

Typical Subspace

Working in the orthogonal basis $\{|x\rangle\}$ in which ρ is diagonal, we may regard our quantum source as an effectively classical source, the eigenvalue playing here the same role as the probability in Shannon theory. The quantum source generates messages that are strings $(|\mathbf{x}\rangle = |x_1\rangle|x_2\rangle|x_3\rangle\dots|x_n\rangle)$ of ρ 's eigenstates $(|x\rangle)$, each with a probability

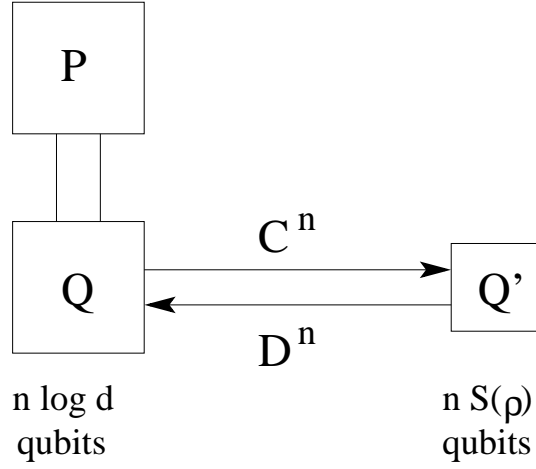


Figure 6.4: The compression operation C^n compresses a quantum source Q stored in $n \log d$ qubits into $n S(\rho)$ qubits, while preserving the entanglement with the reference system P . The source is accurately recovered via the decompression operation D^n without damaging the initial entanglement with P .

given by the product of the corresponding eigenvalues. Therefore, it makes sense to talk of an ϵ -typical sequence, x_1, \dots, x_n for which

$$\left| \frac{1}{n} \log \frac{1}{\lambda_{x_1} \lambda_{x_2} \dots \lambda_{x_n}} - S(\rho) \right| \leq \epsilon. \quad (6.22)$$

in exactly the same way as for classical typical sequences. An ϵ -typical state is a state $|x_1\rangle|x_2\rangle\dots|x_n\rangle$ for which the sequence x_1, x_2, \dots, x_n is ϵ -typical. Define the ϵ -typical subspace to be the space spanned by all ϵ -typical states. We will denote the ϵ -typical subspace by $T(n, \epsilon)$, and the projector onto the ϵ -typical subspace by $P(n, \epsilon)$,

$$P(n, \epsilon) = \sum_{\epsilon\text{-typical}} |x_1\rangle\langle x_1| \otimes |x_2\rangle\langle x_2| \otimes \dots \otimes |x_n\rangle\langle x_n|. \quad (6.23)$$

The theorem of typical sequences can be generalized to the quantum case:

Theorem of typical subspace

1. $\rho^{\otimes n}$ is ϵ -typical: Fix $\epsilon > 0$. Then for any $\delta > 0$, for sufficiently large n , the probability that the source outputs an ϵ -typical state is at least $1 - \delta$:

$$\text{Tr}(P(n, \epsilon) \rho^{\otimes n}) \geq 1 - \delta. \quad (6.24)$$

2. Size of the typical subspace: For any fixed $\epsilon > 0$ and $\delta > 0$, for sufficiently large n , the dimension $|T(n, \epsilon)| = \text{Tr}(P(n, \epsilon))$ of $T(n, \epsilon)$ satisfies:

$$(1 - \delta) 2^{n(S(\rho) - \epsilon)} \leq |T(n, \epsilon)| \leq 2^{n(S(\rho) + \epsilon)}. \quad (6.25)$$

3. Small subspaces have zero probability: Let $S(n)$ be a projector onto any subspace of $\mathcal{H}^{\otimes n}$ of dimension at most 2^{nR} , where $R < S(\rho)$ is fixed. Then for any $\delta > 0$, and for sufficiently large n ,

$$\text{Tr}(S(n) \rho^{\otimes n}) \leq \delta. \quad (6.26)$$

The proof of points (1) and (2) results directly from the classical version of the theorem. (3) uses the property of the operators $P(n, \epsilon)$ and $S(n)$ to show that for large n the dimension of $S(n)$ is exponentially smaller than $T(n, \epsilon)$, see [136, 170] for more details on (3).

Noiseless Channel Coding theorem

Using the typical subspace theorem it is not difficult to prove the quantum analogue of Shannon's noiseless channel coding theorem. The main difference with the classical case is that the operation must preserve the quantum coherence of the system.

Theorem 8 (Schumacher's noiseless channel coding theorem) *Let $\{p(x), |x\rangle\}$ be an i.i.d. quantum source, where $\rho = \sum_x p(x)|x\rangle\langle x|$. If $R > S(\rho)$ then there exists a reliable compression scheme of rate R for the source. If $R < S(\rho)$ then any compression scheme of rate R is not reliable.*

The proof is a direct generalization of Shannon's version and is strikingly similar to the classical one. The encoding is done by first measuring whether the output string lies in the typical subspace or not (using the orthogonal projectors $P(n, \epsilon)$ and $I - P(n, \epsilon)$). If the state lies in the typical subspace nothing more is done. If it does not, then replace the state of the system with some standard state $|0\rangle$ chosen from the typical subspace. It follows that the encoding is a map $\mathcal{C}^n : \mathcal{H}^{\otimes n} \rightarrow \mathcal{H}_c^n$ into the 2^{nR} -dimensional subspace \mathcal{H}_c^n , with operator-sum representation

$$\mathcal{C}^n(\sigma) = P(n, \epsilon)\sigma P(n, \epsilon) + \sum_i A_i \sigma A_i^\dagger, \quad (6.27)$$

where $A_i = |0\rangle\langle x_i|$, $|x_i\rangle$ is an orthogonal basis for the subspace orthogonal to the typical subspace $T(n, \epsilon)$ and $|0\rangle$ is a blank state. The decoding operation $\mathcal{D}^n : \mathcal{H}_c^n \rightarrow \mathcal{H}^{\otimes n}$ is defined to be the identity on \mathcal{H}_c^n . It is then easy to prove $R < S(\rho)$ using point (1) of the theorem of typical subspaces. The converse is proven using point (3), similarly as for Shannon theorem, see [112, 136, 170].

Practical Data compression

Classical encoding techniques such as Huffman coding [51] can be adapted to quantum sources with three constraints:

1. The encoding must be completely reversible (unitary).
2. The encoding must preserve the quantum coherence (superpositions) of the states. Which is equivalent to preserving the entanglement between systems Q and its reference P .
3. The encoding must erase the original state in the process of creating the compressed one, as we have to fulfill the no-cloning theorem.

6.4 Quantum and Classical Correlations

In this section we show that the von Neumann entropy has also an interpretation as the measure of the amount of quantum correlations $[qq]$ (e-bits) and classical bits $[cc]$ that can be extracted from a pure bipartite states $|\psi\rangle_{AB}$.

Entanglement Distillation of Pure States

Given a pure bipartite state $|\psi\rangle_{AB}$,

$$|\psi\rangle_{AB} = \sum_x \sqrt{p(x)} |x\rangle_A |x\rangle_B, \quad (6.28)$$

with partial trace $\text{Tr}_{A(B)}[|\psi\rangle\langle\psi|_{AB}] = \rho$. Entanglement distillation transforms n copies of $|\psi\rangle_{AB}$, using local operations ($\mathcal{E}_A \otimes \mathcal{E}_B$), into nR Bell pairs $|\Phi^+\rangle$ ($|\Phi^+\rangle = [|00\rangle + |11\rangle]/\sqrt{2}$), where R is the rate of the protocol. The nR Bell pairs being a maximally entangled state, they can be written as,

$$|\Phi^+\rangle_{AB}^{\otimes nR} = \sum_{i=1}^{2^{nR}} \sqrt{\frac{1}{2^{nR}}} |i\rangle_A |i\rangle_B. \quad (6.29)$$

The n -fold tensor product $|\psi\rangle_{AB}^{\otimes n}$ reads,

$$|\psi\rangle_{AB}^{\otimes n} = \sum_{x_1, x_2, \dots, x_n} \sqrt{p(x_1)p(x_2)\dots p(x_n)} |\mathbf{x}\rangle_A |\mathbf{x}\rangle_B. \quad (6.30)$$

where $\mathbf{x} = x_1 x_2 \dots x_n$. The entanglement distillation of pure states is strikingly similar to the distillation of classical correlations from a noiseless probability distribution. Here we need to construct a bijection assigning an $|i\rangle$ to each $|x_1 x_2 \dots x_n\rangle$ preserving the quantum coherence and being implementable by local operations and communication. It is easy to see that the optimal solution is to assign to each i a typical sequence of \mathbf{x} which gives the bound $R < S(\rho)$. Having $2^{nR} > 2^{nS(\rho)}$ will oblige us either to assign some i to a non-typical sequence \mathbf{x} or to implement mappings, such as $|x\rangle|x\rangle \rightarrow [|ii\rangle + |i'i'\rangle]/\sqrt{2}$, which are impossible by local operations as they generate entanglement.

Achievability

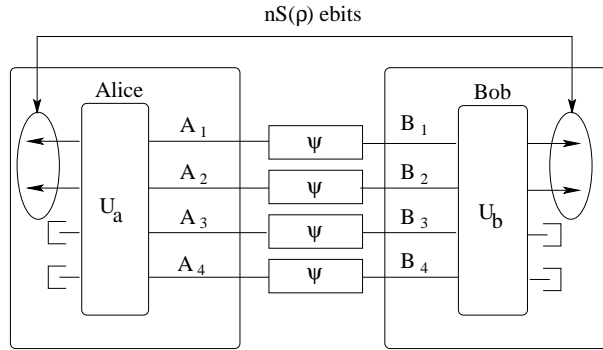


Figure 6.5: Applying Schumacher compression on each side to n copies of a bipartite pure state we obtain $2^{nS(\rho)}$ Bell pairs $|\Phi^+\rangle$.

Once we understand that the compression operation \mathcal{C}^n preserves the entanglement between the system Q and the reference P it is easy to generalize it to the distillation of entangled bipartite pure states, by applying a projection on the typical subspace on both locations A and B , as shown in Fig. 6.5. We obtain a maximally entangled state (all the Schmidt coefficients are equal) of rank $2^{nH(\lambda_x)} = 2^{nS(\rho)}$, which is equivalent to $nS(\rho)$ ebits.

Classical Correlations

By a similar reasoning as the one used for entanglement distillation one can prove that the maximal amount of classical correlations that can be extracted from the n -fold tensor product $|\psi\rangle^{\otimes n}$ are $nS(\rho)$ bits, which can be reached by applying a projective measurement over the Schmidt basis at each side.

6.5 State Merging

Consider an i.i.d. source which distributes a n product bipartite states ρ_{AB} between Alice (A) and Bob (B). In order to simplify the discussion we assign a reference system E that purifies the states ρ_{AB} , as shown in Fig 6.6, where E can be seen as the environment.

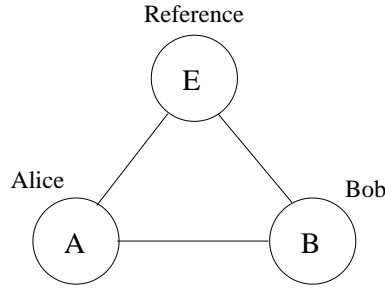


Figure 6.6: An i.i.d. source distributes a n product bipartite states ρ_{AB} between Alice (A) and Bob (B). We assign a reference system E that purifies the states ρ_{AB} .

Similarly to the classical "data merging", at the end of the quantum "state merging" protocol Bob has control of the whole joint state ρ_{AB} . There are two important differences with the classical case. Firstly, in the quantum case Alice has to erase her state in the process, as we have to fulfill the no-cloning theorem. Secondly, in some cases ($S(A|B) < 0$) the protocol generates entanglement.

As in the classical case, the quantity related to state merging is the conditional entropy $S(A|B)$, the interpretation being different if the conditional entropy is positive or negative. During the state merging protocol either we need to expend quantum communication ($S(A|B) > 0$) or we are able to generate entanglement ($S(A|B) < 0$). To quantify the entanglement generated we define the coherent information $I(A)B = -S(A|B)$.

Theorem 9 (Data Merging) *Suppose (A, B) is an i.i.d. distributed source shared by Alice (A) and Bob (B). Suppose $R > S(A|B)$, then there exists a reliable protocol that succeeds to transfer A from Alice to Bob either by using on average R qubits of noiseless communication ($S(A|B) > 0$) or distilling $I(A)B$ e-bits of entanglement ($S(A|B) < 0$) on average by using $S(A:E)$ bits of classical communication. Conversely, if $R < S(A|B)$ the data merging will not be reliable.*

The Optimal Rates

The proof of $R > S(A|B)$ is strikingly similar to the classical case. The key idea is that the initial entanglement between two partners sharing a pure bipartite state cannot increase by local operations and classical communication (LOCC) and sending R qubits increases the entanglement at most by R e-bits. We separate the proof in two different cases depending on the sign of $S(A|B)$.

Positive Conditional Entropy

As in the classical case we regroup Alice and the Reference into a single partner, see Fig 6.7. Using the entropy as a measure of entanglement of a pure state (section 6.4)

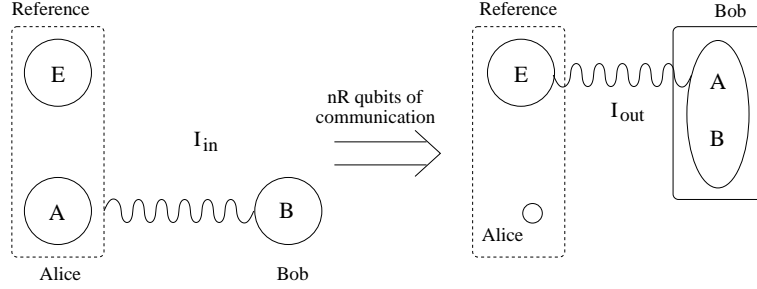


Figure 6.7: Alice sends R qubits of quantum communication to Bob in order to give Bob access to the joint state (A, B) . Alice and the Reference are inside a dashed rectangle to stress that for technical reason of the proof they are considered as a unique partner. At the beginning of the protocol B is entangled with (A, E) where at the end it is only entangled with E .

the initial entanglement between Bob and Alice+Reference reads,

$$E_{in} = S(B). \quad (6.31)$$

After Alice has sent R qubits through a quantum noiseless channel to Bob and successfully implemented the state merging protocol, the final entanglement between Bob and Alice+Reference reads,

$$E_{out} = S(A, B), \quad (6.32)$$

as at the end of the merging Bob has the full state ρ_{AB} . Because communicating R qubits can increase at most the entanglement by R e-bits ², we obtain the bound,

$$E_{in} + R \geq E_{out}, \quad (6.33)$$

which gives $R > S(A|B)$.

Negative Conditional entropy

Here the initial situation is similar, but the protocol changes as it will generate entanglement. As shown in Fig. 6.8, by applying local operations $(\mathcal{E}_A \otimes \mathcal{E}_B)$ after a successful state merging Alice and Bob share R e-bits of entanglement and Bob has on his side the state ρ_{AB} which is entangled with the reference, The final entanglement between Bob and Alice+Reference then reads,

$$E_{out} = S(A, B) + R. \quad (6.34)$$

Because the entanglement cannot increase through local operations and classical communication

$$E_{in} \geq E_{out}, \quad (6.35)$$

we get the lower bound $R \leq I(A|B) = -S(A|B)$.

²Sending R half of Bell pairs through a quantum noiseless channel is the optimal solution.

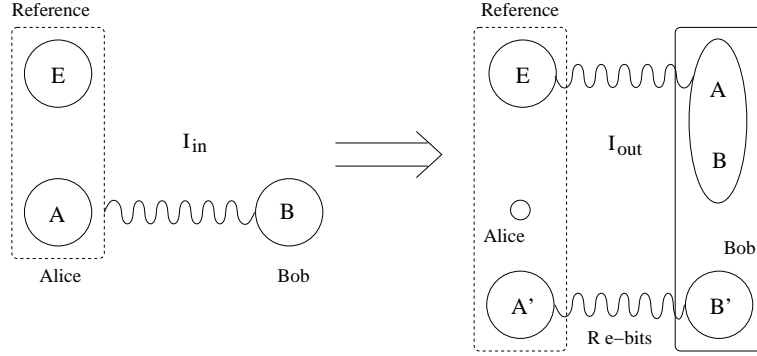


Figure 6.8: At the same time that Alice merges her system A with Bob's system B they succeed to generate R e-bits of entanglement (system (A', B')).

Achievability

The proof works on the asymptotic limit where Alice and Bob share n realizations of the joint distribution: $\rho_{AB}^{\otimes n}$.

Positive Conditional Entropy

The solution for the case $S(A|B) > 0$ is a trivial generalization of the classical case. Alice divides her typical subspace T_A into $2^{nS(A|B)}$ subspaces and maps each subspace into a given single letter message $|i\rangle$. After sending $|i\rangle$, which costs her R qubits on average, Bob can reconstruct $\rho_{AB}^{\otimes n}$ from $\rho_B^{\otimes n}$ and $|i\rangle$. The only subtlety appearing in the quantum case is that the operation must preserve the coherent superpositions and Alice must erase her state in the process.

Negative Conditional Entropy

Let us summarize the proof of [59, 109]. Alice decorrelates her quantum system A from Eve by applying an incomplete measurement that maps $\mathbf{A} = A_1, A_2, \dots, A_n$ into a subspace $\tilde{\mathbf{A}}$ of dimension $2^{nI(A)B}$ with measurement outcome j of rank $nS(A:E)$. After Alice decorrelation operation, given the outcome j , the tripartite state reads $|\Psi^j\rangle_{\tilde{\mathbf{A}}\mathbf{B}\mathbf{E}}$. Where Alice and Eve system reads,

$$\rho_{\tilde{\mathbf{A}}\mathbf{E}} = \tau_{\tilde{\mathbf{A}}} \otimes \rho_{\mathbf{E}}, \quad (6.36)$$

where $\tau_{\tilde{\mathbf{A}}}$ is a maximally mixed state of dimension 2^{nR} . Then Alice communicates j to Bob using $nS(A:E)$ bits of classical communication. Because Bob holds the purification of $\rho_{\tilde{\mathbf{A}}\mathbf{E}}$ (he controls modes $\mathbf{B} = B_1, B_2, \dots, B_n$ and knows the result j of Alice's measurement) and all the purifications are equivalent up to a local isometry (see appendix A), there is a $U_{\mathbf{B}}^j : \mathbf{B} \rightarrow \tilde{\mathbf{B}}\mathbf{B}$ such that

$$(\mathbb{I}_{\tilde{\mathbf{A}}\mathbf{E}} \otimes U_{\mathbf{B}}^j) |\Psi^j\rangle_{\tilde{\mathbf{A}}\mathbf{B}\mathbf{E}} = |\Phi^+\rangle_{\tilde{\mathbf{A}}\mathbf{B}}^{\otimes nR} \otimes |\Psi\rangle_{\tilde{\mathbf{B}}\mathbf{B}\mathbf{E}}, \quad (6.37)$$

as $|\Phi^+\rangle_{\tilde{\mathbf{A}}\mathbf{B}}$ is a valid purification of $\tau_{\tilde{\mathbf{A}}}$ and $|\Psi\rangle_{\tilde{\mathbf{B}}\mathbf{B}}$ of $\rho_{\mathbf{E}}$. We observe that the merging protocol succeeds to send Alice's system to Bob at the same time that we distill $nI(A)B$ ebits. In the process Alice had to communicate $nS(A:E)$ bits to Bob in order to tell Bob which $U_{\mathbf{B}}^j$ to apply. The protocol can then be summarized using the following resource inequality,

$$S(A:E)[c \rightarrow c] + \{qq\} \geq I(A)B[qq]. \quad (6.38)$$

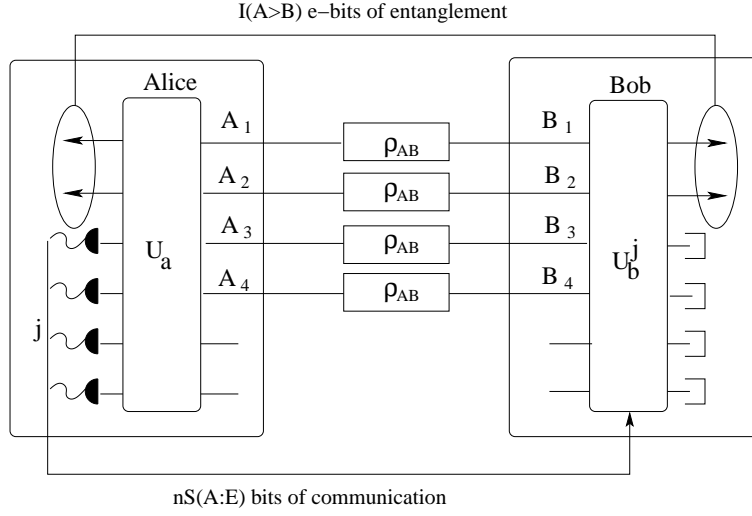


Figure 6.9: Alice applies a partial measurement that maps \mathbf{A} into $\tilde{\mathbf{A}}$ (dimension $nI(A)B$) and outputs j (rank $nS(A:E)$). Then Alice communicates j to Bob who applies a unitary operation U_B^j that allows distill him to nR ebits.

A Family of Protocols

As shown in [1], the previous entanglement distillation protocol can be made fully quantum by replacing the communication of $nS(A:E)$ classical bits by sending $nS(A:E)/2$ qubits through a quantum noiseless channel. Alice applies a Schumacher compression

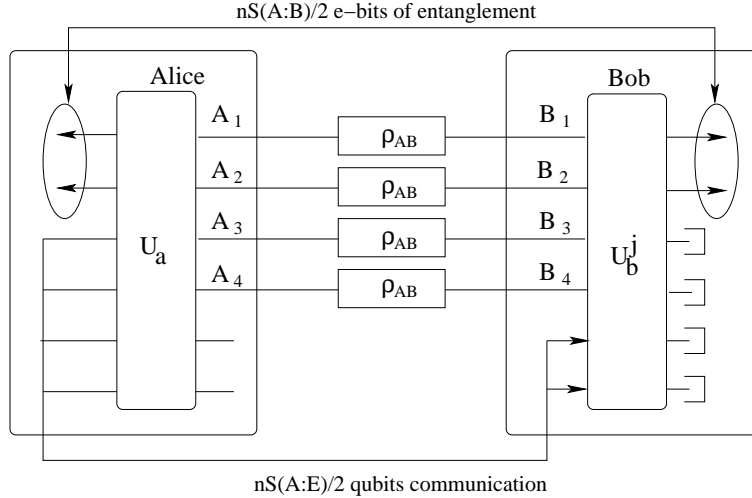


Figure 6.10: Alice applies a Schumacher compression on her system \mathbf{A} , subsequently applies a given unitary operation U_A , and factors the output in two subsystems \mathbf{A}_1 and \mathbf{A}_2 of size $nR_1 = S(A:B)/2$ and $nR_2 = S(A:E)/2$ qubits, respectively. She sends \mathbf{A}_2 to Bob who applies a local isometry $U_{\mathbf{A}_2\mathbf{B}} : \mathbf{A}_2\mathbf{B} \rightarrow \mathbf{B}_1\hat{\mathbf{B}}\mathbf{B}$ distilling in the process nR ebits.

on her system \mathbf{A} (size $nS(A)$ qubits) and subsequently applies a given unitary opera-

tion U_A ³ and factors the output in two subsystems \mathbf{A}_1 and \mathbf{A}_2 of size nR_1 and nR_2 qubits, respectively. By sending \mathbf{A}_2 to Bob Alice completely decorrelates \mathbf{A}_1 from \mathbf{E} , if U_a is chosen properly. After Bob receives \mathbf{A}_2 he has a purification of the system $\rho_{\mathbf{A}_1} \otimes \rho_{\mathbf{E}}$, then by applying a local isometry $U_{\mathbf{A}_2\mathbf{B}} : \mathbf{A}_2\mathbf{B} \rightarrow \mathbf{B}_1\hat{\mathbf{B}}\mathbf{B}$ he can transform the tripartite system into

$$(\mathbb{I}_{\mathbf{A}_1\mathbf{E}} \otimes U_{\mathbf{A}_2\mathbf{B}})|\Psi\rangle_{\mathbf{ABE}} = |\Phi^+\rangle_{\mathbf{A}_1\mathbf{B}_1}^{\otimes nR_1} \otimes |\Psi\rangle_{\hat{\mathbf{B}}\mathbf{B}\mathbf{E}}. \quad (6.39)$$

Because sending nR_2 qubits of communication cannot increase the entanglement by more than nR_2 e-bits, and nR_1 being the number of distilled e-bits at the end of the protocol, using similar techniques as before we get the inequality,

$$R_2 + S(B) \geq S(A, B) + R_1. \quad (6.40)$$

Because the sizes of \mathbf{A}_1 and \mathbf{A}_2 correspond to \mathbf{A} we have the constraint,

$$R_1 + R_2 = S(A). \quad (6.41)$$

Combining equations (6.40) and (6.41) we obtain $R_1 \leq S(A:B)/2$ and $R_2 \geq S(A:E)/2$, which gives the following resource inequality,

$$\frac{1}{2}S(A:E)[q \rightarrow q] + \{qq\} \geq \frac{1}{2}S(A:B)[qq]. \quad (6.42)$$

This is called the *mother* resource inequality and generates other known quantum protocols by appending (a) or prepending (p) either the teleportation (TP) protocol [21]

$$2[c \rightarrow c] + [qq] \geq [q \rightarrow q], \quad (6.43)$$

or the dense coding (DC) protocols [22]

$$[q \rightarrow q] + [qq] \geq 2[c \rightarrow c], \quad (6.44)$$

as shown in [58]. The three *children* of the mother protocol are, a noisy version of teleportation

$$S(A:B)[c \rightarrow c] + \{qq\} \geq I(A:B)[q \rightarrow q] \quad : (\text{a} - \text{TP}), \quad (6.45)$$

a noisy version of dense coding

$$S(A)[q \rightarrow q] + \{qq\} \geq S(A:B)[c \rightarrow c] \quad : (\text{a} - \text{DC}), \quad (6.46)$$

and one-way entanglement distillation,

$$S(A:E)[c \rightarrow c] + \{qq\} \geq I(A:B)[qq] \quad : (\text{p} - \text{TP}), \quad (6.47)$$

which is exactly the protocol presented before.

Quantum Channel Capacity

Previously we were considering a static scenario where a bipartite source distributed a given quantum state ρ_{AB} to Alice and Bob and subsequently they applied entanglement distillation to extract e-bits. Here we are going to study a dynamic scenario where the source ($|\psi\rangle_{AA'}$) is located at Alice's site and Alice sends the system A' through the

³For more details on how to construct U_A see [1].

quantum channel \mathcal{N} to Bob. The previous mother protocol can be shown to have a source-channel dual called the *father* protocol

$$\frac{1}{2}S(A:E)[qq] + \{q \rightarrow q\} \geq \frac{1}{2}S(A:B)[q \rightarrow q], \quad (6.48)$$

that can be deduced from the mother protocol as explained in [1]. The father protocol has two children, the entanglement-assisted classical information transmission

$$S(A)[qq] + \{q \rightarrow q\} \geq S(A:B)[c \rightarrow c] \quad : (\text{a} - \text{DC}), \quad (6.49)$$

and the quantum communication protocol,

$$\{q \rightarrow q\} \geq I(A)B[q \rightarrow q], \quad (6.50)$$

which can be found by prepending the entanglement distribution $[q \rightarrow q] \geq [qq]$ ⁴ to the father protocol.

Analogy with Classical Communication In the classical communication through noisy channels ($\{c \rightarrow c\}$), one can get rid of the noiseless channel ($[c \rightarrow c]$) used in the error correction by selecting a proper code \mathcal{C}_l . Similarly, in the quantum communication scenario we can get rid of the $\frac{1}{2}S(A:E)$ ebits by using the protocol described in [59], inspired on the secret key distillation presented in next chapter.

Quantum Capacity

For a given source-channel pair $\{|\psi\rangle_{AA'}, \mathcal{N}\}$, Alice's and Bob's final state reads

$$\rho_{AB} = (\mathbb{I}_A \otimes \mathcal{N}_{A' \rightarrow B})|\psi\rangle_{AA'}, \quad (6.51)$$

The coherent information $I(\mathcal{N}, \rho) = I(A)B$, where $\rho = \text{Tr}_A[|\psi\rangle\langle\psi|_{AA'}]$, gives the amount of entanglement that can be distributed among the channel using i.i.d. sources. Optimizing among the different i.i.d. sources for a fixed channel \mathcal{N} we obtain the i.i.d. quantum capacity

$$C_1 = \max_{\rho} I(\mathcal{N}, \rho). \quad (6.52)$$

One could expect, by analogy with the classical case, that C_1 gives the optimal capacity of the quantum channel. Unfortunately this is not the case in quantum information as shown in [61, 179] for very noisy depolarizing channels and more recently in [185] for general depolarizing channels. This clearly shows that i.i.d. sources are not sufficiently general. The most general definition of the capacity reads,

$$C = \lim_{n \rightarrow \infty} \frac{1}{n} \max_{\rho_n} I(\mathcal{N}^{\otimes n}, \rho_n), \quad (6.53)$$

where we must consider the behavior of the channel on input states entangled across many uses (ρ_n).

6.6 Classical Communication

In this section we will derive how to distribute classical correlations through quantum channels. In order to simplify the discussion we use the entanglement-based description of communication over quantum channels, as shown in Fig. 6.11. After Alice and Bob measurements, the joint state (a, b) being completely classical (diagonal density

⁴Which is trivially implemented by sending an e-bit through a noiseless channel.

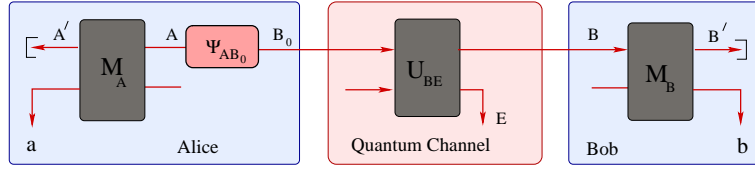


Figure 6.11: Alice source is modelled by combining an entanglement source $|\psi\rangle_{AB_0}$ and a POVM measurement M_A . Alice's message B_0 is sent through the quantum channel, which is modelled by an unitary interaction U_{BE} . Finally Bob applies a POVM measurement M_B on B .

matrix), the correlations between Alice and Bob are given by the mutual entropy $S(a:b)$. Sometimes we will use the notation

$$I(\rho_{AB}; M_A, M_B) = S(a : b), \quad (6.54)$$

to stress that the joint probability distribution (a, b) has been obtained by applying POVMs M_A and M_B on the bipartite state $\rho_{AB} = \mathbb{I} \otimes \mathcal{N}(|\psi\rangle_{AB_0})$.

Accessible Information

If we restrict Bob's measurements to individual (product) measurements ($M_B = M_B^{\otimes n}$) over each message of the i.i.d. source sent by Alice through the quantum channel, as in Fig. 6.12, the optimal mutual information that both partners can reach reads,

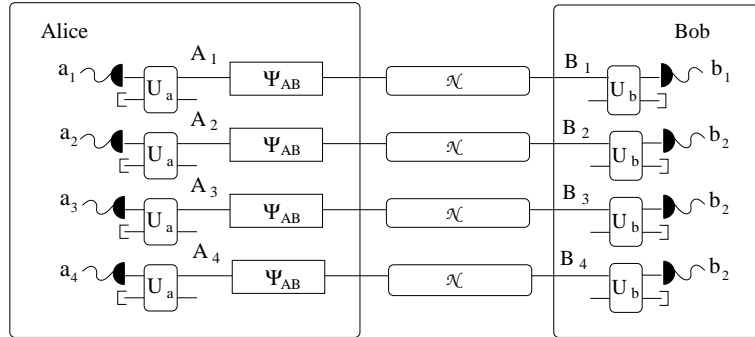


Figure 6.12: Alice generates an i.i.d. source by applying individual POVM measurements over n copies of an entangled state. After sending half of the modes through the quantum channel Bob applies an individual measurement over each signal.

$$I_{acc}(\rho_{AB}; M_A) = \max_{M_B} I(\rho_{AB}; M_A, M_B), \quad (6.55)$$

which is called the *accessible information*. Optimizing over the possible C-Q sources \mathcal{S} (the best combination entangled state+measurement $\mathcal{S} = \{|\psi\rangle_{AB_0}, M_A\}$) we obtain the *product states and product measurement capacity*

$$C_{11}(\mathcal{N}) = \max_{\mathcal{S}} I_{acc}(\mathbb{I} \otimes \mathcal{N}(|\psi\rangle_{AB_0}); M_A). \quad (6.56)$$

The accessible information can be very difficult to calculate as it needs an optimization over all the different possible measurements. Interestingly a simpler quantity called the Holevo bound gives a useful upperbound.

Holevo Bound

For a given source $\{p(a), |\varphi\rangle_{B_0}\}$ and for a fixed measurement M_B on Bob side, Alice and Bob mutual entropy is upperbounded by the so-called *Holevo bound*,

$$S(a : b) \leq S(a:B) = S(\rho_B) - \sum_a p(a) S(\rho_B^a). \quad (6.57)$$

$S(a:B)$ being a function of the quantum state B preceding the measurement, it does not depend on Bob's measurement M_B . The accessible information is then also upperbounded by $S(a:B)$. Interestingly, the calculation of $S(a:B)$ is strikingly simple as it only depends on the von Neumann entropies of Bob's state ρ_B and of Bob's state conditioned on Alice data a (ρ_B^a).

The proof of the Holevo bound can be made strikingly simple by combining the physical model of measurement, entanglement-based description of sources (see appendix A) and the strong subadditivity of von Neumann entropy.

Proof

Consider a quantum bipartite state ρ_{AB} shared by Alice and Bob, as shown in Fig. 6.13. Suppose that after receiving mode B , Bob applies the quantum operation \mathcal{T} .

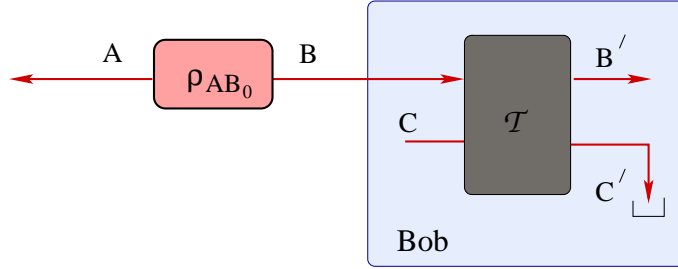


Figure 6.13: Alice and Bob share a quantum state ρ_{AB} . Bob applies a quantum operation \mathcal{T} over mode B which can be modeled by applying a unitary operation U_{BC} on B and an ancillary system C .

Using the physical description of a quantum operation (see appendix A) we can model \mathcal{T} by applying a unitary operation U_{BC} on B and an ancillary system C . Without loss of generality we can consider that the ancilla is initially decorrelated from the bipartite state ($\rho_{ABC} = \rho_{AB} \otimes \rho_C$) giving

$$S(A:B) = S(A:B, C) \quad (6.58)$$

where we used $S(A, C) = S(A) + S(C)$ and $S(A, B, C) = S(A, B) + S(C)$ as C is decorrelated from A and B . Because a unitary interaction does not increase the entropy of a system (the eigenvalues do not change) we have,

$$S(A:B, C) = S(A:B', C') = S(A:B') + \underbrace{S(A:C'|B')}_{\geq 0}. \quad (6.59)$$

After discarding the ancilla C' and using the strong subadditivity of the entropy we finally obtain

$$S(A:B') \leq S(A : B). \quad (6.60)$$

This result being true for any quantum system on Alice side, it holds also when Alice's state is the result of a POVM measurement (ρ_{aB}),

$$S(a:B') \leq S(a : B). \quad (6.61)$$

Measurements M_B being a subclass of the existing quantum operations \mathcal{T} that Bob can apply, as explained in appendix A, they must satisfy equation (6.61), giving the Holevo bound ($B' = b$).

Achieving the Holevo Bound

Unfortunately the accessible information does not generally achieve the Holevo bound. One can see that in order to saturate the Holevo bound using product measurements the states ρ_B^a must have orthogonal support, which is not generally satisfied. Interestingly one can saturate the Holevo bound if we allow Bob to apply collective measurements, which are more general than product measurements.

HSW Codes

The technique used to achieve the Holevo bound is pretty similar to the encoding used in the communication over classical channels.

Correlation Distillation

In order to saturate the Holevo bound Alice divides her set of typical sequences T_a into $2^{nS(a|B)}$ non-overlapping subset \mathcal{C}_l of size $2^{nS(a:B)}$, the so-called HSW codes [104, 171]. Bob assigns a collective POVM measurement M_l to each code \mathcal{C}_l , where each POVM M_l allows him to determine which of the $2^{nS(a:B)}$ sequences of \mathcal{C}_l was really measured by Alice, see [136, 97] for a detailed description of how to construct such a measurement. In order to extract perfect correlations Alice first applies $M_A^{\otimes n}$ to the n modes \mathbf{A} and

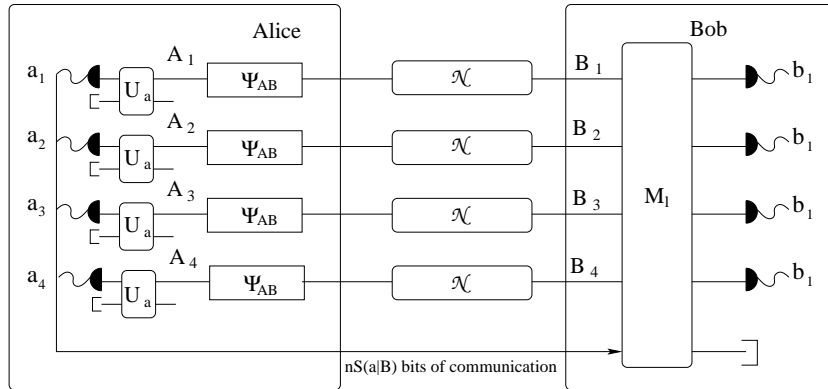


Figure 6.14: Alice i.i.d. C-Q source can be modeled by applying a POVM over each entanglement source. In full generality in order to reach $nS(a:B)$ bits of correlation Bob would have to apply a collective measurement over all his received states.

communicates to Bob in which code \mathcal{C}_l lies her measurement outcome $\mathbf{a} = a_1 a_2 \dots a_n$, see Fig. 6.14. Knowing in which code \mathcal{C}_l lies \mathbf{a} , Bob applies the corresponding POVM (M_l) which allows him to determine \mathbf{a} without error.

Correlation Distribution

Similarly as in the case of quantum communication or Shannon communication, the correlation distillation protocol can be transformed into a correlation distribution without the need of a supporting noiseless channel. To do so, Alice actively selects a given HSW code \mathcal{C}_l among all the typical sequences of T_a . By just sending sequences among this code and Bob applying the corresponding POVM M_l Alice succeeds to send $nS(a:B)$ classical bits of correlation to Bob efficiently.

Classical Communication Capacity

Optimizing over the possible C-Q sources, or equivalently searching the best combination entangled state+measurement $\mathcal{S} = \{|\psi\rangle_{AB_0}, M_A\}$, we obtain the *product states and collective measurement capacity*

$$C_{1\infty}(\mathcal{N}) = \max_{\mathcal{S}} S(a : B). \quad (6.62)$$

Note that by the Holevo bound this is an upperbound of the product states and product measurement capacity C_{11} defined in equation (6.56).

An even more general scenario of correlations distribution allows Alice to use non-i.i.d sources. In this scenario the capacity reads,

$$C(\mathcal{N}) = \lim_{n \rightarrow \infty} \frac{1}{n} C_{11}(\mathcal{N}^{\otimes n}). \quad (6.63)$$

An important open question in quantum information theory is whether the conjecture $C_{1\infty}(\mathcal{N}) = C(\mathcal{N})$ is true. This is called the additivity of the "classical capacity of quantum channels" problem, which is related to other conjectures in quantum information [178]. Up to now, there is only a proof for some channels such as the unital channels [117] and entanglement-breaking channels [177].

6.7 Continuous-Variable Entropy

Any N -mode quantum state of a continuous variable system is completely described by the density operator

$$\rho = \sum_{\mathbf{n}, \mathbf{m}=0}^{\infty} \rho_{\mathbf{n}, \mathbf{m}} |n_1, \dots, n_N\rangle \langle m_1, \dots, m_N|. \quad (6.64)$$

The von Neumann entropy (6.3) of ρ is calculated by calculating the Shannon entropy of the eigenvalues of ρ , as described previously. Contrary to the Shannon entropy of continuous distributions here the entropy is well-defined, as Fock basis is infinite but discrete. As in the case of discrete variables one can define quantum conditional, mutual and relative entropies for continuous variable systems. In the following we will restrict to the case of Gaussian states as it is the only family of states that we use in the following chapters.

Gaussian States

The displacement being a unitary operation and the entropy being invariant under unitary operations ⁵, we have

$$S(\rho) = S(D(\alpha)^\dagger \rho D(\alpha)). \quad (6.65)$$

⁵As it is defined using a trace operation ($S(\rho) = -\text{Tr}[\rho \log \rho]$) and the trace is invariant under unitary operations.

It is then trivial to see that the mean value does not play any role in the calculation of the entropy. We can then restrict ourselves to the family of states with null mean value. As we explained in Chapter 2 every Gaussian state ρ_G with null mean value and covariance matrix γ can be decomposed into a tensor product of thermal states $\lambda_k \mathbb{I}$ by applying a proper symplectic transformation S ,

$$S\gamma S^T = \bigotimes_{k=1}^N \begin{bmatrix} \lambda_k & 0 \\ 0 & \lambda_k \end{bmatrix}, \quad (6.66)$$

where the symplectic eigenvalues are the solution of a polynomial with the symplectic invariants as coefficients, as described in Chapter 2.

Symplectic transformations being a subclass of unitary operations, the von Neumann entropy of ρ_G and $\bigotimes_k \lambda_k \mathbb{I}$ must be equal. This proves that the problem of calculation of the von Neumann entropy of a N -mode Gaussian state reduces to calculating the von Neumann entropy of a product of N thermal states $\lambda_k \mathbb{I}$,

$$S(\rho_G) = \sum_{k=1}^N S(\lambda_k \mathbb{I}). \quad (6.67)$$

Thermal State

The density operator of a thermal state reads,

$$\rho = \sum_{n=0}^{\infty} \frac{\langle n \rangle^n}{(\langle n \rangle + 1)^{n+1}} |n\rangle \langle n|, \quad (6.68)$$

where $\langle n \rangle$ is the mean number of thermal photons in the state. The density matrix being diagonal on the Fock basis, its von Neumann entropy reads

$$\begin{aligned} S(\rho) &= -\frac{1}{\langle n \rangle + 1} \sum_{i=0}^{\infty} \left(\frac{\langle n \rangle}{\langle n \rangle + 1} \right)^i \log \left[\left(\frac{\langle n \rangle}{\langle n \rangle + 1} \right)^i \frac{1}{\langle n \rangle + 1} \right] \\ &= -\frac{1}{\langle n \rangle + 1} \sum_{i=0}^{\infty} \left(\frac{\langle n \rangle}{\langle n \rangle + 1} \right)^i \left[i \log \langle n \rangle - i \log(\langle n \rangle + 1) - \log(\langle n \rangle + 1) \right] \\ &\stackrel{(a)}{=} -\frac{1}{\langle n \rangle + 1} \left[\langle n \rangle (\langle n \rangle + 1) (\log \langle n \rangle - \log(\langle n \rangle + 1)) - (\langle n \rangle + 1) \log(\langle n \rangle + 1) \right] \\ &= (\langle n \rangle + 1) \log(\langle n \rangle + 1) - \langle n \rangle \log \langle n \rangle, \end{aligned} \quad (6.69)$$

where in (a) we used $\sum_{i=0}^{\infty} x^i = 1/(1-x)$ and $\sum_{i=0}^{\infty} ix^i = x/(1-x)^2$.

Having in mind that the relation between the mean photon number and the symplectic eigenvalue of a thermal state is $\lambda = 2\langle n \rangle + 1$ we obtain,

$$S(\lambda_k \mathbb{I}) = G[(\lambda_k - 1)/2], \quad (6.70)$$

where

$$G(x) = (x+1) \log(x+1) - x \log x. \quad (6.71)$$

Chapter 7

Quantum Key Distribution

7.1 Introduction

Armies, governors, diplomats and politicians have always tried to secure their communications from a potential adversary. This has always been accompanied with a willing of intercepting the adversary secret messages, in order to place theirself in an advantageous strategic position. This has generated a continuous demand of improving secret coding methods as well as decoding techniques all along the human history. From the simple methods used during the Antiquity based on simple transpositions of the alphabet to modern methods based on the difficulty of solving some mathematical problems such as factoring large numbers on a usual computer, the history of cryptography has been a continuous competition between cryptographs developping novel coding techniques and others trying to crack them. See reference [184] for a fascinating introduction to the history of cryptography. The development of telecommunications all along the 20th century came along with an increasing demand for treating huge amounts of data. In order to deal with such a quantity of data the encoding and decoding techniques were automatized, first by using mechanized engines and later with computers. During the last decades of the 20th the demand for secure communications has expanded to new domains of activity such as financial and commercial communications, a phenomenon that has been amplified with the development of the Internet.

All the cryptographic techniques from the simplest to the most evolved one have a similar working procedure. Two partners, usually called Alice and Bob by the cryptographic community, want to communicate a message (the plaintext) secretly from a potential eavesdropper, usually called Eve. Alice applies an encoding algorithm having as inputs the plaintext and an encoding key. This turns the plaintext into a cyphertext non-understandable to Eve. Then Bob decodes the cyphertext sent by Alice using a decoding algorithm with the help of a second key. Alice and Bob keys do not need to be the same, as in asymmetric cryptography used in public encryption techniques such as RSA [160], which is the encoding algorithm most commonly used nowadays. In RSA the encoding key (Alice) is public, allowing anyone to encode a message and send it to Bob. On the other hand the decoding key is only known by Bob, so he is the only one able to decode the message.

Unconditional Security

All the encryption techniques developed in the past have been sooner or later broken, usually requiring the effort of large groups of brilliant people working intensively during long periods of time. The cracking of the Enigma encoding machine of the German

army during the Second World War by the Allies scientists working at Bletchley Park [184] is such an example. Modern cryptographic techniques do not escape to this problem. Most of them being based on the difficulty of solving quickly enough some mathematical problems, such as factorizing large numbers in RSA [160], there is no guarantee that one day we will found an algorithm that solves those problems in an efficient time.

One can rise the following question: "does it exist an encoding technique that allows for a completely unbreakable secure communication?". The surprising answer is yes, the technique is called the One-Time Pad, where the key is the same length as the plaintext. If the key is secure, truly random, and never reused, the cipher is unbreakable. The protocols is extremely simple as show in Fig. 7.1: Alice applies a modular addition to the plaintext p^n and the secret key k^n ($c^n = p^n \oplus k^n$) and sends the cyphertext c^n to Bob. After receiving the string c^n Bob undoes the modular addition $p^n = c^n \oplus k^n$, recovering the plaintext p^n . The idea of the protocol is to transform the plaintext into a completely noisy message non-understandable by Eve by applying random noise. In order to decode the message Bob needs to know exactly the noise added by Alice, the secret key.

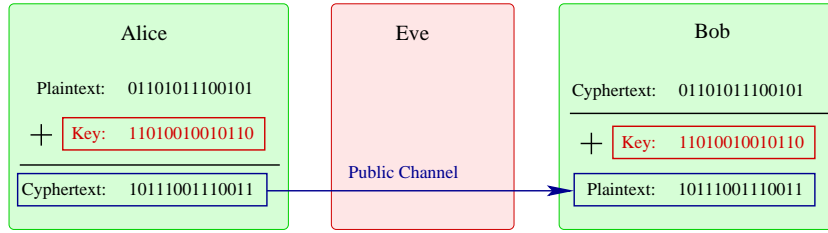


Figure 7.1: Alice applies a modular addition to the plaintext p^n and the secret key k^n ($c^n = p^n \oplus k^n$) and sends the cyphertext c^n to Bob. After receiving the string c^n Bob undoes the modular addition $p^n = c^n \oplus k^n$, recovering the plaintext p^n .

In 1949 Shannon [176], fixed the conditions to have "perfect secrecy" using his recently developed information theory [175]: Alice encodes the plaintext X using a key K and an encoding algorithm $C = f(X, K)$ that outputs the cyphertext C that she later sends to Bob. In order to have perfect security we need that Eve does not gain any information by knowing the cyphertext sent by Alice. This translates in the Information Theory language into the condition

$$H(X|C) = H(X), \quad (7.1)$$

where we demand that the uncertainty on the plaintext knowing the cyphertext being equal to the initial uncertainty on the plaintext. The conditional entropy $H(X|C)$ can be rewritten as

$$H(X|C) = H(X) + H(C|X) - H(C). \quad (7.2)$$

In the case of One-Time Pad, the modular addition $c^n = x^n \oplus k^n$ implies $H(C|X) = H(K|X)$, the plaintext and the key been independent, we finally obtain $H(C|X) = H(K)$. By construction of the protocol we have $H(C|K) = H(X|K) = H(X)$, which together with the concavity of entropy ($H(C|K) \leq H(C)$) gives $H(C) \geq H(X)$. Finally, in order to satisfy the condition (7.1) we need the secret key entropy being larger than the plaintext amount of information,

$$H(K) \geq H(X). \quad (7.3)$$

The best way of implementing One-Time Pad is Alice applying compression to his message up to $nH(X)$ bits and then using $nH(X)$ bits of secret key to generate the

cyphertext by applying a modular addition to the plaintext and the secret key. In [176] Shannon proved that all the protocols achieving "perfect security" are equivalent to the One-Time Pad.

The Problem of Distributing the Secret Key

If we inspect the One-Time Pad protocol we see that the "perfect security" is based on a cornerstone assumption, "Alice and Bob initially share a secret key". In practice it is impossible to have a guaranteed perfectly secure secret key distribution by classical cryptographic techniques, we have to trust some messenger or communication channel that distributes the secret keys between Alice and Bob. Despite that, One-time Pad has been widely used, usually combined with a public cryptographic protocol that generates the secret key.

Surprisingly, at the middle of the 80's Bennett and Brassard [20] showed that the secret key distribution can be made "perfectly" secure if one uses Quantum Key Distribution (QKD) techniques. The security of QKD is based on the no-cloning theorem [199] that shows that it is impossible to perfectly copy ensembles of non-orthogonal quantum states. Any attack by a potential adversary being detectable by Alice and Bob as it would disturb the quantum communication signal.

7.2 Classical Key Distribution

Before presenting the theory of Quantum Key Distribution we are going to analyze a classical analogue, allowing us to present some of the key elements that we will later generalize in QKD. We will show how two partners can distill secret bits from a bipartite distribution correlated with the potential eavesdropper Eve by using public communication over a classical channel.

Definition of Resources We will use the notation $[ss]$ for the static resource corresponding to a *secret bit* shared between Alice and Bob, and $\{ss\}$ will denote a bipartite correlated distribution shared by Alice, Bob and the potential eavesdropper Eve. The dynamical resource $[s \rightarrow s]$ corresponds to the distribution of a secret bit between Alice and Bob, and $\{s \rightarrow s\}$ to a eavesdropped channel.

Static Scenario

Consider three parties Alice, Bob and the eavesdropper Eve sharing a tripartite distribution (X, Y, Z) generated by an independent and identically distributed (i.i.d.) source with distribution $p(x, y, z)$, as shown in Fig. 7.2. In order to extract a secret key (unknown by Eve) from their shared data, Alice and Bob have to apply secret key distillation. In what follows we will concentrate on *one-way distillation* protocols where one of the trustful partner's data (Alice or Bob) is used as reference to generate the key and the communication travels on one single direction.

Secret Key Distillation Helped by a Private Channel

In order to simplify the proof let's first consider that Alice and Bob have access to prior secret keys and a private channel. A private key being a perfectly correlated sequence, the first thing that Alice and Bob have to do is to distill a noiseless correlated sequence.

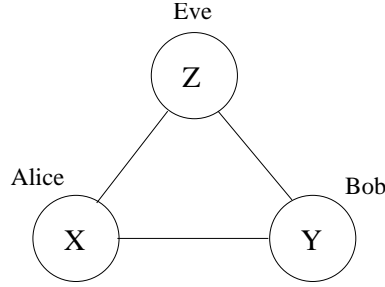
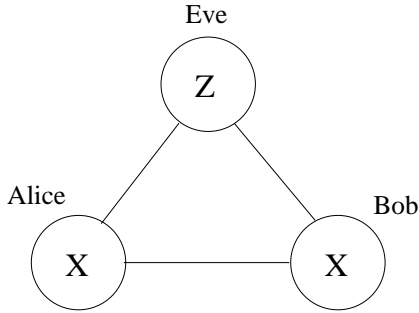


Figure 7.2: Alice, Bob and the eavesdropper Eve sharing a tripartite distribution (X, Y, Z) generated by an i.i.d. source.

a) After Error Correction



b) After Privacy Amplification

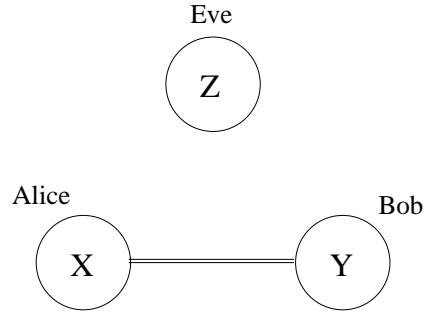


Figure 7.3: a) After the error correction Alice and Bob data are perfectly correlated, but remain correlated to Eve. b) Privacy amplification allows Alice and Bob to completely decorrelate their joint data from Eve.

Error Correction To do so they can decide to fix Alice data as the reference and apply data merging as explained in Chapter 5, by Alice communicating $nH(X|Y)$ bits to Bob through a private channel ($nH(X|Y)[s \rightarrow s]$). After the error correction Alice and Bob sequences are the same, as shown in Fig. 7.3, generating $nH(X)$ bits of correlation.

Privacy Amplification After the error correction Alice and Bob data remains correlated to Eve data (Z). Alice or Bob could then individually decorrelate their own data from Eve by using $nH(X:Z)$ random bits as explained in section 5.6, but this would decrease their shared correlations. To overcome this problem Alice and Bob use $nH(X:Z)$ shared secret bits (denoted $nH(X:Z)[ss]$) as input of their decorrelation protocol. This allows them to decorrelate their data from Eve, as by definition she has no information on the secret bits and at the same time it preserves their joint correlations as both partners apply the same deterministic operation.

Secret Key Rate At the end of the protocol Alice and Bob share a perfectly correlated distribution (X, X) unknown by Eve which generates $nH(X)$ bits of secret key ($nH(X)[ss]$). The protocol can be resumed using the following resource inequality,

$$H(X|Y)[s \rightarrow s] + H(X:Z)[ss] + \{ss\} \geq H(X)[ss]. \quad (7.4)$$

The net gain of secret key reads,

$$K = H(X) - H(X|Y) - H(X:Z) = H(X:Y) - H(X:Z), \quad (7.5)$$

which corresponds with the well known result of Csiszar and Korner [52]. A symmetric counterpart exists if we interchange Alice and Bob roles, Bob data Y being now the reference,

$$K = H(X:Y) - H(Y:Z). \quad (7.6)$$

Secret Key Distillation

Following the results of section 5.5, Alice divides her typical set T_X into $2^{nH(X|Y)}$ codes C_l of size $2^{nH(X:Y)}$. When the secret key rate K is positive ($K = H(X:Y) - H(X:Z) >$

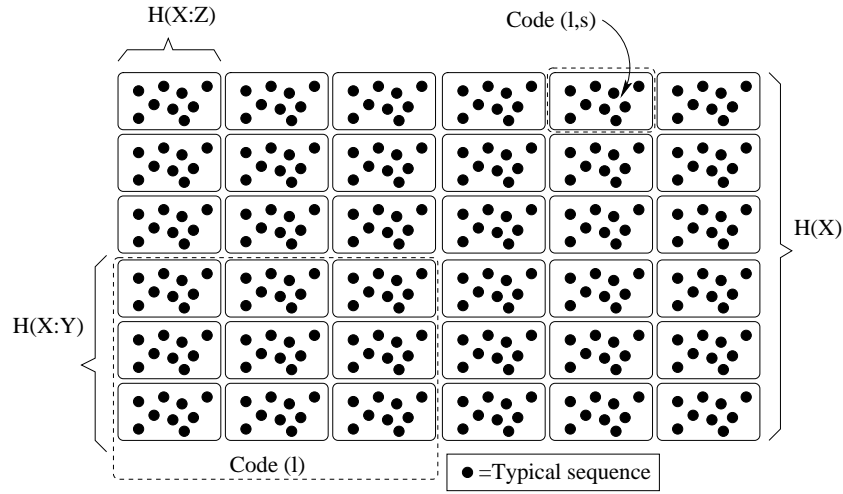


Figure 7.4: Alice divides the Typical set T_X into $2^{nH(X|Y)}$ codes C_l of size $2^{nH(X:Y)}$ labeled by l (error correcting codes). Each code C_l is again divided into $2^{nK} = 2^{n[H(X:Y) - S(X:Z)]}$ privacy amplification codes $C_{l,s}$ of size $2^{nH(X:Z)}$ labeled (l, s) , where s encodes the secret key.

0), each code C_l can be divided into 2^{nK} codes $C_{l,s}$ of size $2^{nS(X:Z)}$, as shown in Fig. 7.4. Eve in order to estimate x^n needs Alice revealing in which of the $2^{nH(X|Z)}$ codes $C_{l,s}$ lies x^n . Alice by revealing just l' during the error correcting allows Bob to perfectly estimate x^n but leaves Eve on the uncertainty of knowing on which of the 2^{nK} codes $C_{l',s}$ lies x^n . The secret key being just the information on s which is known by Alice and Bob but unknown by Eve, generating nK secret bits.

One can then get rid of the prior secret key and the use of a private channel without error using the preceding technique that can be resumed in the following resource inequality,

$$H(X|Y)[c \rightarrow c] + \{ss\} \geq (H(X:Y) - H(X:Z))[ss], \quad (7.7)$$

where now the communication can be done over a public channel without compromising the security of the key, as in Csiszar and Korner work [52]. If the error correction is done in the opposite direction (from Bob to Alice) one can derive the symmetric relation

$$H(Y|X)[c \rightarrow c] + \{ss\} \geq (H(X:Y) - H(Y:Z))[ss]. \quad (7.8)$$

The big problem of classical key distribution is that no physical law forbids Eve from monitoring the communication of the channel and getting the same information as

Bob receives, which gives a null secret key. Then in order to have a positive secret key we need to do some assumptions on Eve capability to extract information. Hopefully, Quantum Key Distribution (QKD) techniques make possible the distribution of a secret key because of the physical constraint imposed on Eve by the no-cloning theorem [199].

Secret Key Distribution

One can transform the previous static protocol of secret key distillation into a secret key protocol if Alice sends a copy of the data generated by her source X to Bob through a noisy channel, Bob receives a noisy version Y of X and Eve extracts information Z by monitoring the channel, as shown in Fig. 7.5. Subsequently Alice and Bob apply

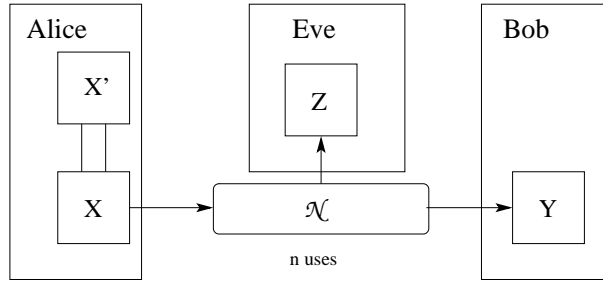


Figure 7.5: Alice sends a copy of the data generated by her source X to Bob through a noisy channel. Bob receives a noisy version Y of X and Eve extracts information Z by monitoring the channel.

error correction and privacy amplification to their data in order to extract a secret key.

7.3 Quantum Key Distribution

A Quantum Key Distribution protocol [92] is divided in two steps; Firstly, a quantum communication part where Alice prepares and sends quantum signals through a quantum channel to Bob, who measures them. This is usually implemented over optical fibers in order to benefit from the huge speed and low decoherence of light. Secondly, by running a classical post-processing protocol through a classical authenticated channel Alice and Bob extract a secret key from their correlated data. In the following we will describe a general QKD protocol for binary alphabets that can be easily generalized to continuous variables.

Quantum Communication Step

Before presenting a general prepare-and-measure QKD protocol, we are going to introduce a simple example, the first proposed quantum key distribution protocol, called BB84 [20], where the bits are encoded on a given internal degree of freedom of the photon, such as the polarization. An alternative encoding system more suited to fiber transmission (usual telecom fibers does not preserve the polarization) is the so called *time-bin encoding* [92], where the bits are encoded in the time delay of the photon plus a phase. The encoding is achieved by using an interferometer with arms of different lengths combined with a controlled phase shift on one of both arms.

BB84

In BB84 Alice first generates two random bits for each photon that she will send to Bob. The first bit (x_i) will be used to generate the secret key and the second (b_i) is used to choose which basis Alice would use to encode the bit x_i , either the computational basis $\{|0\rangle, |1\rangle\}$ or the conjugate basis $\{|+\rangle, |-\rangle\}$. The state sent to Bob depends on both bits x_i and b_i such that $\{|0\rangle, |-\rangle\}$ encodes $x_i = 0$ and $\{|1\rangle, |+\rangle\}$ encodes $x_i = 1$. Finally, Alice sends the quantum state to Bob who measures either the computational or conjugate basis depending on the value of his own random bit b'_i , obtaining the result y_i .

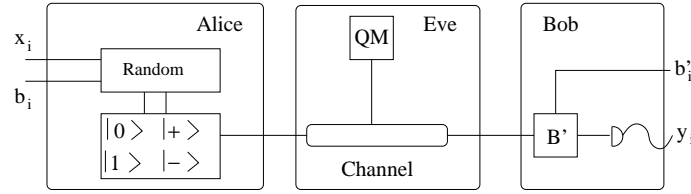


Figure 7.6: Alice generates two random bits that defines which state among $\{|0\rangle, |1\rangle, |+\rangle, |-\rangle\}$ she sends to Bob through the quantum channel. Bob randomly chooses which basis (b') he is going to measure. The state sent by Alice decohered after the interaction with Eve ancilla which is stored in a quantum memory (QM).

The random switching between two conjugate bases is crucial in order to secure the protocol against intercept-resend attacks by Eve. If Alice was encoding always on the same basis Eve could just measure the right basis and then re-prepare and send the same state to Bob. Fortunately, the no-cloning theorem forbids perfectly cloning two non-orthogonal states (ex: $\{|0\rangle$ and $|-\rangle\}$ encoding the same bit $x_i = 0$), making any intercept-resend attack detectable if Alice and Bob change between conjugate bases. The protocol is built in such a way that Eve having no information on b_i , the optimal attack she can apply is to interact with the signal sent by Alice in order to make an imperfect clone and keep it on a quantum memory until Alice reveals the chosen basis to Bob.

Generalization

A general protocol starts by Alice choosing randomly among l different ensembles $\mathcal{E}_l = \{p_{i,l}, |\psi_{i,l}\rangle\langle\psi_{i,l}|\}$, where $p_{i,l}$ is the probability of sending $|\psi_{i,l}\rangle\langle\psi_{i,l}|$ once we have chosen the ensemble l , with all the ensembles \mathcal{E}_l giving the same average state σ ,

$$\forall l : \sum_i p_{i,l} |\psi_{i,l}\rangle\langle\psi_{i,l}| = \sigma, \quad (7.9)$$

in order to forbid Eve from knowing which basis ensemble was chosen. Subsequently Alice randomly chooses which state $|\psi_{i,l}\rangle\langle\psi_{i,l}|$ among the ensemble \mathcal{E}_l she will send to Bob through the quantum channel. The choice of l corresponds to the basis (b_i in BB84) and i to the data that will be used later to generate the secret key (x_i in BB84). In an entanglement-based description of the protocol Alice starts from a pure entangled state $|\Psi\rangle_{AB_0}$, applies a randomly selected POVM measurement \mathcal{M}_l which generates the ensemble \mathcal{E}_l , as explained in Chapter 7, and subsequently sends the system B_0 to Bob through the quantum channel. Once the signal arrives to Bob station, he applies a randomly selected POVM measurement $\mathcal{N}_{l'}$, where each \mathcal{N}_k is constructed in order to optimize the secret key for a given preparation \mathcal{E}_k of Alice.

Most of the QKD protocols, such as BB84, are constructed in such a way that all the ensembles \mathcal{E}_l are a resolution of the identity $\sigma = \mathbb{I}$ (BB84: $|0\rangle\langle 0| + |1\rangle\langle 1| = |+\rangle\langle +| + |-\rangle\langle -| = \mathbb{I}$) and Alice and Bob POVM measurements use the same projective measurement ($\mathcal{N}_{l'} = \mathcal{M}_l = \sum_i |\varphi_i^l\rangle\langle \varphi_i^l|$) on the entanglement-based description. In the prepare-and-measure scheme, this is equivalent to Bob measuring the orthogonal basis which correspond with the ensembles of states sent by Alice. Nearly all of the existing protocols use either conjugate or mutually unbiased bases, as in BB84, which give no correlations between Alice and Bob when they use different bases.

Classical Post-Processing Step

Once Alice and Bob have collected a sufficiently large list of correlated data, they proceed with the classical post-processing step. They first apply *sifting*, where they compare Alice encoding basis l and Bob measurement basis l' and keep only those data where they have used the same basis. Subsequently they apply *parameter estimation*, where by revealing a randomly chosen sample of their data they obtain an estimation of the parameters of the channel, which allows them to upperbound Eve information I_E . Now they are ready to extract a secret key from the remaining data by using *error correction* and *privacy amplification* [190], using similar techniques as those used in Section 7.2. The size of the secret key reads,

$$K = I_C - I_E - M, \quad (7.10)$$

where I_C is the amount of correlations that Alice and Bob could extract over a noiseless channel, I_E is an upperbound of Eve information and M is the size of the message disclosed during the error correction. Hence, it is important that the error correction discloses as little information as possible and that the parameter estimation upperbounds correctly Eve's information.

Choice of the Bases

Alice and Bob choosing their bases randomly implies the loss of a fraction of the data during the sifting step of the classical post-processing, due to a mismatch of the selected bases. As an example, in the usual version of BB84 Alice and Bob randomly select among two conjugate bases using a balanced probability distribution ($p(b_1) = p(b_2) = 1/2$) which forces Alice and Bob to discard half of their data. Interestingly, an equiprobable distribution between the bases is not compulsory, one can reduce the amount of data loss by using an unbalanced distribution between bases b_1 and b_2 . Consider the case where we use basis b_1 most of the time (probability $P(b_1) = 1 - p$, p small) and the conjugate basis b_2 is used just to estimate the channel, which in the entanglement-based description corresponds to a tomographic reconstruction of ρ_{AB} ¹. Now the fraction of lost data is reduced to $2p(1 - p)$ and $1 - 2p$ of the data is used to extract the secret key, where $2p^2$ of the data is used for tomographic reconstruction of the channel.

In what follows we will consider that Alice and Bob use the optimal single measurement in order to extract a secret key and the rest of measurements are just used to estimate the channel, it is then enough to analyze the tripartite state corresponding to a single measurement,

$$\rho_{\mathbf{aBE}} = \sum_{a^n} P(a^n) |a^n\rangle\langle a^n|_{\mathbf{a}} \otimes \rho_{\mathbf{BE}}^{a^n}. \quad (7.11)$$

¹Fortunately, usually a fully tomographic reconstruction is not needed, as selecting a proper set of measuring bases, as in BB84, is enough to have an upperbound of Eve information.

From $\rho_{\mathbf{aBE}}$ one can calculate the secret key rate K , which must be corrected by a prefactor $s = 1 - 2p$ ($K' = sK$) taking into account the sifting procedure. The price to pay when we want to increase the prefactor s is that we have to wait for longer times in order to collect enough statistics to estimate the channel.

One can also avoid losing data ($s = 1$) if Bob has access to a quantum memory. Bob just needs to store the quantum signal sent by Alice in a quantum memory until she reveals the selected ensemble \mathcal{E}_l , after which Bob applies the correct measurement. Unfortunately, even if quantum memories have been implemented experimentally in the lab, they are extremely hard to implement, making this alternative unrealistic for a commercial implementation.

7.4 Security Against Eavesdropping

In quantum key distribution, in order to have unconditional security we evaluate the secret key rate by upper bounding the information that the adversary (Eve) can acquire in the worst case scenario. This is typically done under the following assumptions:

1. Eve has no limit in terms of computational power.
2. Eve has full control over the quantum channel, and is only limited in her action on this channel by the laws of quantum physics.
3. Eve can freely monitor the classical public channel used for key distillation, but she cannot modify the messages (authenticated channel) ².
4. Eve has no access to the laboratories (apparatuses) of Alice and Bob.

In order to be unconditionally secure, a QKD protocol must be secure against an attack where Eve is allowed to prepare any global ancillary system and make interact it collectively with all the pulses sent by Alice, as shown in Fig. 7.7. After the communication of n pulses Alice-Bob-Eve tripartite state reads,

$$\rho_{\mathbf{aBE}} = \sum_{a^n} P(a^n) |a^n\rangle \langle a^n|_{\mathbf{a}} \otimes \rho_{\mathbf{BE}}^{a^n}. \quad (7.12)$$

where $\rho_{\mathbf{aBE}}$ is not necessarily an i.i.d. state $\rho_{\mathbf{aBE}}^{\otimes n}$ as the different pulses sent by Alice can be entangled by Eve's interaction.

After having monitored the public communication between Alice and Bob during the classical post-processing, Eve applies the optimal joint measurement over all her ancilla, as shown in Fig. 7.7. The security with respect to this attack, called *coherent attack*, is very complex to address. Fortunately, it was proven for discrete-variable QKD in [155] that under the assumption of the symmetry of the privacy amplification and channel estimation protocols, coherent attacks are not more efficient than collective attacks.

Coherent Attacks are not More Efficient than Collective Attacks In [155] it is shown that assuming the channel estimation and the privacy amplification algorithms being built in a symmetric way, which is generally the case, one can show that collective attacks are as efficient as coherent attacks. This extremely important result shows that under some trivial constraint on the post-processing protocol one can restrict the proof of unconditional security to the class of collective attacks. Interestingly, in collective

²In order to authenticate the channel Alice and Bob need to spend a fixed amount of secret key per round of the QKD protocol, which is obtained by extracting part of the generated secret key. Obviously Alice and Bob need a preexisting key provided by the QKD company (which they have to trust) for their first protocol round.

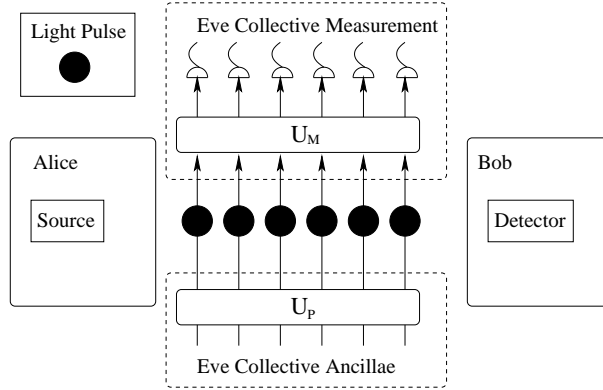


Figure 7.7: Eve prepares her ancilla in a global state by applying a joint unitary operation U_P to the ensemble of all ancilla. After the interaction of the ancilla with the pulses sent by Alice, the ancilla are stored in a quantum memory. Once the classical post-processing is finished Eve applies the optimal global measurement over the ensemble of ancilla, by implementing a joint unitary U_M plus individual measurements.

attacks the tripartite state can be considered being i.i.d. ($\rho_{aBE}^{\otimes n}$), which simplifies very much the calculations, as we can use entropic quantities.

Consider the bipartite ρ_{AB}^N shared between Alice and Bob after Alice sending N pulses to Bob. In [155, 157] the authors show that assuming that ρ_{AB}^N is symmetric, after tracing a small fraction of the modes ($r \ll N$) the state ρ_{AB}^{N-r} can be approximated by a mixture of i.i.d. states,

$$\rho_{AB}^{N-r} \approx \int d\sigma \sigma_{AB}^{\otimes (N-r)} \quad (7.13)$$

with exponentially high precision (on r). Parameter estimation will tell Alice and Bob which of the different σ_{AB} they actually share. One can even relax the symmetric constraint on ρ_{AB}^N if the privacy amplification is built in a symmetric way [156].

Unfortunately the result of [155] only works for discrete variables. This results can not be trivially extended to continuous variable systems, due to some technicalities of taking the limit $d \rightarrow \infty$ in the proof. But this problem also arises when we try to trivially generalize some other information theory proofs to continuous variables, such as the continuity of the von Neumann entropy, where the problem is solved by adding a bound on the energy of the system. This leads us to conjecture the equivalence between coherent and collective attacks for continuous variables in [84].

Collective Attacks

In a collective attack (see Fig. 7.8) Eve prepares her ancillary system in a product state and each ancilla interacts individually with a single pulse sent by Alice, being later stored in a quantum memory. The tripartite state then reads,

$$\rho_{aBE} = \left[\sum_a P(a) |a\rangle \langle a|_a \otimes \psi_{BE}^a \right]^{\otimes n}. \quad (7.14)$$

After listening to the public communication between Alice and Bob during the classical post-processing, Eve applies the optimal collective measurement on the ensemble of stored ancilla. Alice-Bob-Eve tripartite state being i.i.d., the security analysis of collective attacks is much more tractable as we can use entropic quantities.

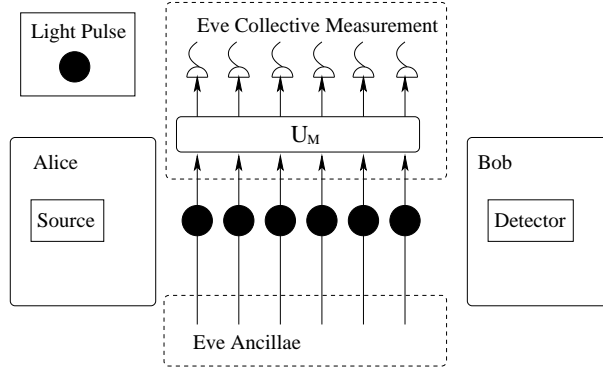


Figure 7.8: Eve ancilla are prepared in a product state. After interaction of each ancilla with a single pulse sent by Alice, the ancilla are stored in a quantum memory. Once the classical post-processing is finished, she applies a global measurement over the ensemble of ancilla, applying a joint unitary U_M plus individual measurement.

Individual Attacks

If Eve has no access to collective measurements, the optimal attack she can perform is the so called individual attack, see Fig. 7.9. In an individual attack Eve interacts

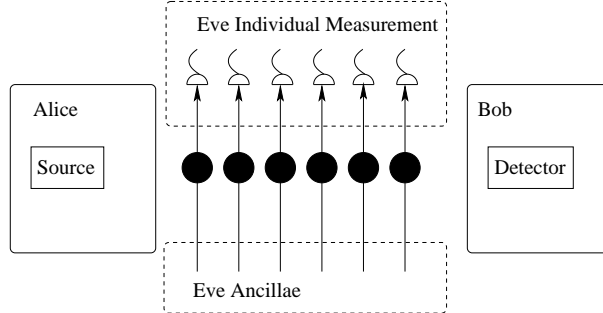


Figure 7.9: Eve ancilla are prepared in a product state. After interaction of each ancilla with a single pulse sent by Alice, the ancilla are stored in a quantum memory. After the sifting step of the post-processing Eve applies individual measurement over her ancilla.

individually with each pulse sent by Alice and stores each ancilla in a different quantum memory. Contrary to preceding attacks Eve now performs an individual measurement on each ancilla, just after the sifting procedure but before the classical post-processing. The tripartite state then reads,

$$\rho_{\mathbf{abe}} = \left[\sum_{x,y,z} P(x,y,z) |x,y,z\rangle \langle x,y,z| \right]^{\otimes n}, \quad (7.15)$$

where z are Eve POVM measurement results. During the post-processing Eve is limited to operations on her classical data.

For a given protocol, the upperbound on Eve information is calculated by running an optimization among all the possible POVM measurements that Eve can apply. For some protocols implemented over highly symmetric channels, as BB84 over a depolarizing channel, Eve individual attacks are as efficient as collective attacks. But this is

not generally the case, for example for Gaussian protocols over Gaussian channels the individual attacks perform worse than collective attacks as we will see in Chapters 8 and 9 of this dissertation.

7.5 One-Way QKD Protocols

In this section we are going to present the generalization of the classical protocols presented in Section 7.2. This family of protocols are called "One-way protocols" as the error correction communication through the public authenticated channel is done in one single direction. We will use an entanglement-based description of the protocols, similarly as in Chapter 7, in order to simplify the analysis and to stress the similarities between entanglement distribution and quantum key distribution. One can see that an e-bit (denoted $[qq]$) is a stronger resource than a secret bit (denoted $[ss]$)

$$[qq] \geq [ss], \quad (7.16)$$

as a secret bit ($[ss]$) can be extracted from an e-bit ($[qq]$) if Alice and Bob measure their respective system over the Schmidt basis (see Appendix A), where no e-bit can be obtained from secret bits. In what follows we will restrict our study to collective attacks by Eve as it is highly probable that they are sufficient to prove unconditional security.

Secret Key from Entanglement Distillation

Consider the situation presented in Fig. 7.10 where Alice generates n entangled states $|\Psi\rangle_{AB_0}$ and sends all systems B_0 through a noisy quantum channel \mathcal{N} . After the

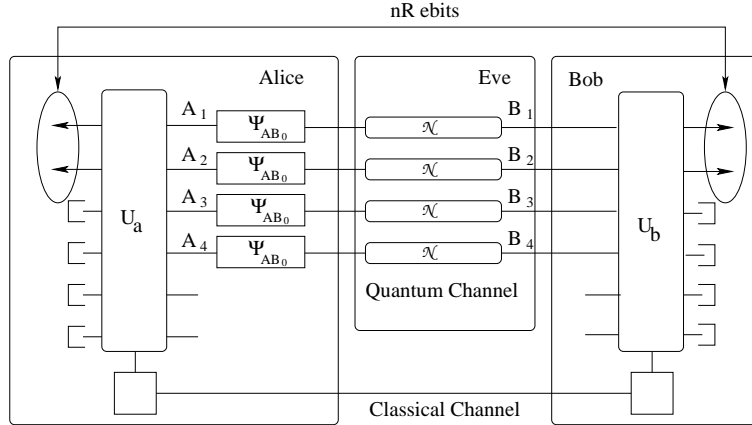


Figure 7.10: Alice distributes n entangled states $|\Psi\rangle_{AB_0}$ through a noisy quantum channel \mathcal{N} . After the distribution Alice and Bob share n copies of a bipartite noisy entangled state ρ_{AB} , which they subsequently transform into nR e-bits, by applying local operations and classical communication (LOCC).

distribution Alice and Bob share n copies of a bipartite noisy entangled state ρ_{AB} , which they subsequently transform into nR e-bits, by applying local operations and classical communication (LOCC).

In the preceding chapter we presented the "state merging protocol" which allows to distill entanglement using local operation and one-way classical communication (1-

LOCC) [108, 109], achieving a rate

$$R_{DR} = I(A|B) = -S(A|B) = S(B) - S(A, B), \quad (7.17)$$

where the "DR" means "Direct Reconciliation", as classical communication is done in the same direction as the quantum communication. Alternatively to this protocol, there is another protocol based on classical communication going from Bob to Alice that achieves a rate

$$R_{RR} = I(B|A) = -S(B|A) = S(A) - S(A, B), \quad (7.18)$$

where the "RR" means "Reverse Reconciliation", as classical communication is done in the opposite direction as the quantum communication. Once Alice and Bob have extracted nR e-bits they could use them to extract nR secret bit by measuring each pair of ebits on the same basis. We have then a way of obtaining a secret key rate of $I(A|B)$ or $I(B|A)$ bits depending on the direction of the error correction.

In many discrete variable channels such as the depolarizing channel both rates (DR and RR) are equal, which is not generally the case. For example, Gaussian protocols over Gaussian channels give different rates in direct and reverse reconciliation.

QKD with Collective Measurement

If we replace the collective operation on Alice side by individual POVM measurements at each mode A_i , as shown in Fig. 7.11, we obtain the usual entanglement-based description of Alice source in an usual prepare-and-measure QKD scheme. The difference with the usual QKD scheme is that Bob instead of measuring independently each pulse, he applies a collective measurement over his n modes $\mathbf{B} = B_1, B_2, \dots, B_n$ in order to reach the Holevo bound $S(a:B)$. One can generalize the classical key distribution

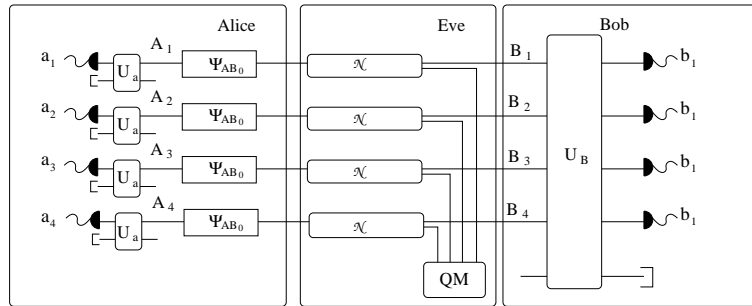


Figure 7.11: If Alice applies individual measurements over each entangled pair, the scheme becomes an entanglement-based description of Alice preparation of a QKD protocol. Bob applies joint measurements over all his modes in order to reach the Holevo bound.

presented in Section 7.2 using the results of Chapter 7.

Error Correction Similarly as in Section 7.2 Alice can use $nS(a|B)$ bits of private communication ($nS(a|B)[s \rightarrow s]$) to communicate Bob on which HSW code [104, 171] \mathcal{C}_l lies a^n , which tells Bob which POVM measurement M_l he has to apply in order to optimally estimate a^n .

Privacy Amplification Using $nS(a:E)$ secret bits ($nS(a:E)[ss]$) Alice and Bob can decorrelate their perfectly correlated data from Eve quantum system, in a very similar way as in the classical case (see [93] for more details), obtaining $nS(a)$ secret bits ($nS(a)[ss]$).

Secret Key Rate One then obtains the following resource inequality,

$$S(a:E)[ss] + S(a|B)[s \rightarrow s] + \{qq\} \geq S(a)[ss]. \quad (7.19)$$

As in the classical case one can get rid of the prior secret bits and private communication, by dividing each HSW code \mathcal{C}_l into 2^{nK} privacy amplification codes \mathcal{C}_{ls} , with

$$K = S(a:B) - S(a:E), \quad (7.20)$$

obtaining the following resource inequality,

$$S(a|B)[c \rightarrow c] + \{qq\} \geq K[ss]. \quad (7.21)$$

Need of Quantum Memories The main difference with classical key distribution is that in order to reach the Holevo bound Bob has to store his quantum state ρ_B until Alice communicates which POVM M_l (which code \mathcal{C}_l) Bob has to use. In full generality, in order to implement collective measurements we need quantum memories, which are very challenging to implement experimentally for one single pulse and impossible for huge amount of data as used in an usual QKD protocol.

Interestingly, for some important channels where all the ρ_B^a (ρ_B^a) can be diagonalized on the same basis [148] the use of quantum memories and collective measurements are not necessary to achieve the Holevo bound as direct individual measurements are enough. Unfortunately this is not the case for continuous-variable Gaussian channels when using homodyne measurements, as even squeezed states and coherent states are non-orthogonal states.

Projective Measurement It is easy to show that in the case of projective measurements on Alice side the previous Direct Reconciliation protocol gives the same rate as the one obtained via entanglement distillation $I(A)B$. We can rewrite equation (7.20) as

$$K = [S(B) - S(B|a)] - [S(E) - S(E|a)]. \quad (7.22)$$

Because Alice-Bob-Eve state is pure we have $S(E) = S(AB)$ and after Alice projective measurement Bob-Eve state Ψ_{BE}^a remains pure implying that $S(B|a) = S(E|a)$, which gives,

$$K = S(B) - S(AB) = I(A)B. \quad (7.23)$$

That is not longer the case when we consider general POVM measurements on Alice side, as for example when Alice applies a noisy measurement.

QKD with Collective Source

The preceding entanglement-based scheme has a symmetric protocol where Bob applies individual measurements and Alice applies a collective measurement. In full generality, Alice source generates entangled states among the different pulses, due to the collective measurement. But after averaging among all the possible outgoing states we obtain an i.i.d. state $\rho_{B_0} \otimes \rho_{B_0} \otimes \dots \otimes \rho_{B_0}$, as result from tracing modes A_i from an i.i.d. state $|\Psi\rangle_{AB_0}^{\otimes n}$. By similar techniques as used previously we can show that there is a one-way protocol that achieves a secret key rate,

$$K = S(A:b) - S(b:E), \quad (7.24)$$

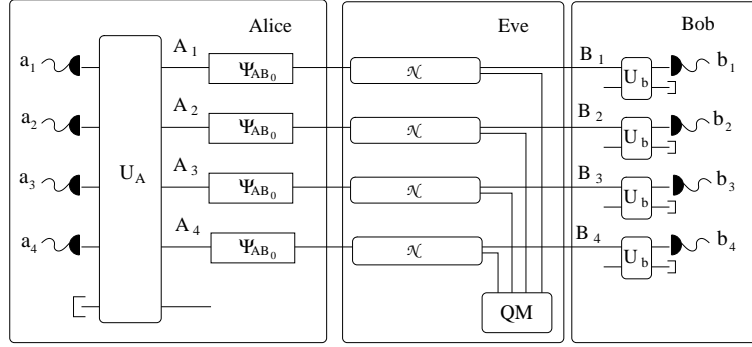


Figure 7.12: An entanglement-based scheme where Alice applies a collective measurement over modes \mathbf{B} is equivalent to a prepare-and-measure source sending entangled pulses. Even so, averaging over all the codewords we obtain an i.i.d. state $\rho_{B_0} \otimes \rho_{B_0} \otimes \dots \otimes \rho_{B_0}$.

where the public communication goes now from Bob to Alice ($H(b|A)$ bits). Similarly as in its symmetric counterpart, if Bob applies a projective measurement the secret key is the same as the coherent information $K = S(A:b) - S(b:E) = I(B)A$.

Realistic QKD

If we now impose a fixed individual POVM measurement on both sides, as shown in Fig 7.13, we obtain the entanglement-based description of an usual QKD protocol, similar to those implemented experimentally.

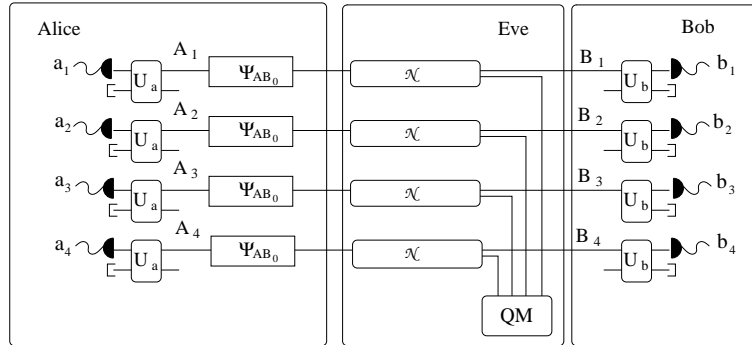


Figure 7.13: Alice and Bob apply individual measurements over all their entangled pairs, which is exactly the previously presented entanglement-based description of QKD.

Similarly as in classical key distribution there is a Direct Reconciliation protocol that achieves a secret key rate,

$$K_{DR} = S(a:b) - S(a:E), \quad (7.25)$$

where the public communication goes from Alice to Bob ($S(a|b)$ bits). There is also a Reverse Reconciliation protocol that achieves a secret key rate,

$$K_{RR} = S(a:b) - S(b:E), \quad (7.26)$$

where the public communication goes now from Bob to Alice ($S(a|b)$ or $S(b|a)$). Notice that in this case the quantum mutual information $S(a:b)$ simply reduces to the Shannon mutual information $I(a:b)$ and the quantum conditional entropies $S(b|a)$ ($S(a|b)$) reduces to the Shannon conditional entropies $H(b|a)$ ($H(a|b)$). The advantage of this protocol is that no quantum memory is needed and that all the post-processing is done on the classical data resulting from the measurements outputs.

QKD with Eve Individual Attacks

Historically there has been an interest on individual attacks by Eve, even if it is known that in most of the cases they do not guarantee unconditional security. In individual attacks Eve is limited to individual measurements over her ancilla and classical post-processing of her measurement output data, as shown in Fig. 7.14. In the case of

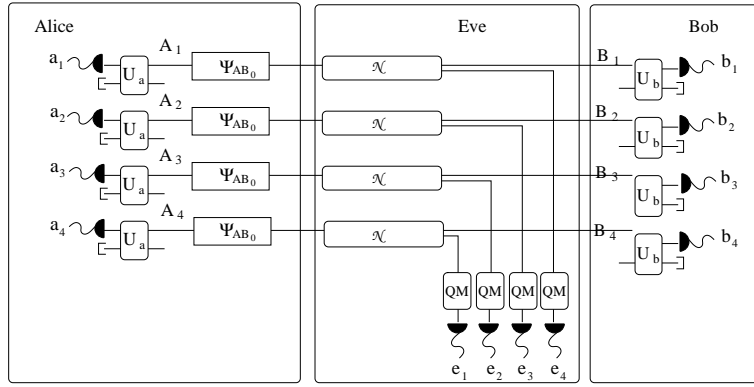


Figure 7.14: In an individual attack Eve has no longer access to a collective measurement over all her ancilla.

direct reconciliation protocols Eve information on Alice data is upperbounded by the accessible information,

$$I_{acc}(\rho_{AE}; M_A) = \max_{M_E} I(\rho_{AE}; M_A, M_E). \quad (7.27)$$

and the secret key rate reads,

$$K_{DR} = I(\rho_{AB}; M_A, M_B) - I_{acc}(\rho_{AE}; M_A). \quad (7.28)$$

In reverse reconciliation protocol the secret key rate reads,

$$K = I(\rho_{AB}; M_A, M_B) - I_{acc}(\rho_{BE}; M_B). \quad (7.29)$$

7.6 Continuous-Variable Quantum Key Distribution

Since the demonstration of continuous-variable quantum teleportation in 1998 [83], there has been a growing interest into the field of continuous variable quantum information, continuous-variable quantum key distribution (CV-QKD) being probably its most fruitful application. Continuous-variable QKD has mainly three practical advantages over qubit implementations such as BB84:

1. In continuous variable quantum processing the repetition rate of the homodyne detection ($\approx 1\text{GHz}$) is much higher than that of the avalanche photodiodes with single photon sensitivity ($\approx 100\text{kHz}$) used in qubit-based QKD.

2. Homodyne detection is far more efficient than single photodetectors, achieving detection efficiencies higher than 90% where single photon detectors currently reach 30%.
3. The class of states generally used in continuous variable QKD, the so called Gaussian states, is easier to generate, being the usual states generated on a quantum optics laboratory.

In the following we are going to present the class of *one-way Gaussian* protocols, where the states sent by Alice are Gaussian mixtures of Gaussian states, Bob applies Gaussian measurements and the error correction communication travels in a single direction.

Squeezed States Protocol

The first idea that one can consider when trying to find a protocol for continuous variables is to generalize the most well know discrete QKD protocol, BB84, described in Section 7.3. Alice sending randomly chosen states of the computational basis $\{|0\rangle, |1\rangle\}$ or the conjugate basis $\{|+\rangle, |-\rangle\}$ translates in continuous variables into Alice generating eigenstates of the x quadrature $\{|x\rangle\}$ or the p quadrature $\{|p\rangle\}$. Bob measurement becoming a balanced homodyne measurement of x or p . Unfortunately the states $\{|x\rangle\}$ and $\{|p\rangle\}$ are infinitely squeezed states which cannot be generated experimentally, as they require sources of infinite energy. Fortunately one can overcome this problem by building a protocol based on finitely squeezed states.

Realistic Squeezed States Protocol The protocols proposed in [44] is based on Alice preparing x -squeezed states displaces along x or p -squeezed states displaces along p , both giving a thermal state of variance V as average output state, as shown in Fig. 7.15. If Alice starts from an x -squeezed vacuum state with covariance matrix,

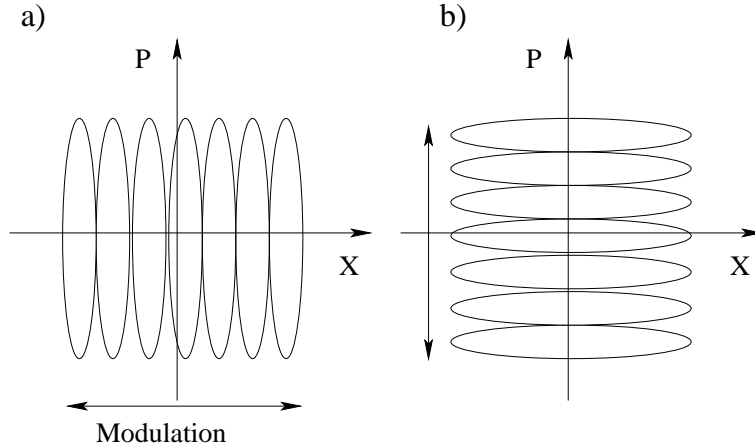


Figure 7.15: a) Alice generates x -squeezed vacuum states (squeezing $1/V$) and displaces them according to a Gaussian distribution (variance $V_A = V - 1/V$). The mixture is equivalent to a thermal state of variance V .

$$\gamma_0 = \begin{pmatrix} e^{-2r} & 0 \\ 0 & e^{2r} \end{pmatrix},$$

where $r \geq 1$ is the squeezing parameter. Then she encodes a random Gaussian-distributed variable a (centered on zero and with variance V_A) into the x -displacement

applied to the squeezed vacuum state ($d : (0; 0) \rightarrow (a, 0)$), as shown in Fig. 7.15 a). Averaging over all possible realizations we get the mixed Gaussian state with null mean value and covariance matrix

$$\gamma_f = \begin{pmatrix} e^{-2r} + V_A & 0 \\ 0 & e^{2r} \end{pmatrix}.$$

We observe that by imposing $e^{-2r} + V_A = e^{2r}$ we obtain a thermal state of variance $V = e^{2r}$. This thermal state is indistinguishable from a thermal state realized by a mixture of p -squeezed states (squeezing parameter r) with Gaussian-distributed p -displacement (variance V_A)

$$\gamma_f = \begin{pmatrix} e^{2r} & 0 \\ 0 & e^{-2r} + V_A \end{pmatrix} = \begin{pmatrix} V & 0 \\ 0 & V \end{pmatrix},$$

as shown in Fig. 7.15 b). As in BB84 we have encoded information in two conjugate quadratures with both output mixed states being the same thermal state of variance V , being then indistinguishable.

The Protocol The quantum communication part of the protocol consists in repeating the following steps for each pulse sent by Alice:

1. Alice generates a random real number (a) from a Gaussian distribution of variance V_A ($V_A = e^{2r}$) and a random bit (b) from a equiprobable binary distribution. At the same time Bob generates a random bit (b').
2. Depending on the value of the random bit (b) Alice sends a x -squeezed state with first moment $d = (a, 0)$ or a p -squeezed state with first moment $d = (0, a)$, where the squeezing r satisfies $V_A = 2 \sinh 2r$.
3. Bob, depending on his random bit (b'), measures either x or p .

After Bob has measured all the pulses, the two partners proceed with the post-processing, which starts by applying sifting:

1. Alice discloses for each pulse the value of b (whether she displaced x or p).
2. Bob keeps only the cases where he measured the right quadrature ($b = b'$).

Finally Alice and Bob apply a *reconciliation* protocols, such as those described in [190], being a combination of error correction and *discretization*. The reconciliation protocols is followed by a *privacy amplification* protocol that extracts the secret key using a given hashing function, see [190] for more details on the post-processing.

Protocol Implementation

As shown in Fig. 7.16 the source (Alice) is based on a master laser beam which is used to generate the *local oscillator* (phase reference) and to pump second harmonic in a nonlinear crystal (SHG). After spectral filtering (F_1) the second harmonic beam pumps an optical parametric amplifier (OPA) which generates a squeezed vacuum state. After filtering the second harmonic (F_2) the squeezed vacuum state is displaced by an amount a by mixing it in a high transmittance beamsplitter (BS_{HT}) with an attenuated coherent state coming from the local oscillator (LO). The attenuation (A) is variable and is a function of a distributed among a Gaussian $G(a)$. After the displacement Alice applies a random $\pi/2$ phase to the local oscillator depending on the value of the bit b (displacement on x or p). Then the quantum signal and the local oscillator travel through the same fiber to Bob multiplexed (M) in time. At Bob station the two signals are demultiplexed (M'), after what Bob applies a random $\pi/2$ phase to the local oscillator depending on the value of the bit b' . Finally Bob applies an homodyne measurement that will measure x or p depending on the value of b' .

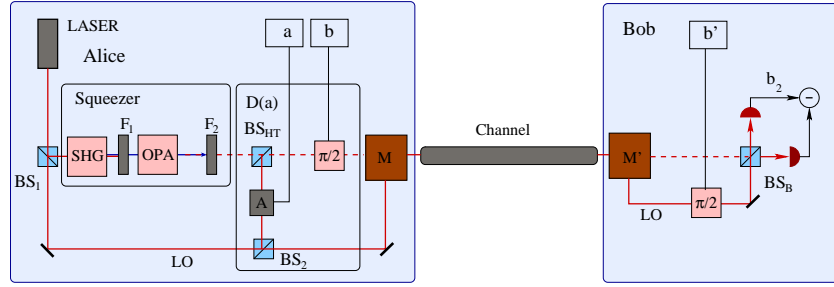


Figure 7.16: Scheme of the optical implementation of the squeezed states protocol.

Coherent States Protocol

A very important step in continuous-variable QKD was the development of a protocol based on Gaussian modulation of coherent states [96], as generating coherent states is far more simpler than squeezed states. The idea is that a thermal state of variance V can also be obtained by a bi-variate Gaussian mixture of coherent states. Alice encodes

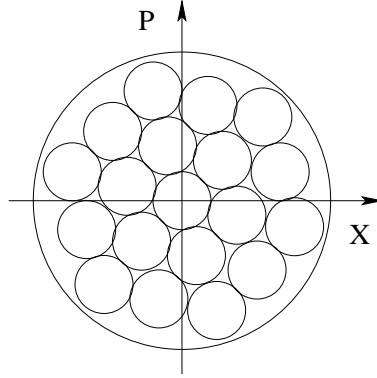


Figure 7.17: a) Alice generates coherent states with random mean value (a_x, a_p) according to a Gaussian distribution (variance V_A). The mixture is equivalent to a thermal state if $V = V_A - 1$.

a random bi-variate Gaussian-distributed variable (a_x, a_p) (centered on zero and with variance V_A) into the (x, p) -displacement applied to the vacuum:

$$\gamma_0 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \longrightarrow \gamma_f = \begin{pmatrix} V_A + 1 & 0 \\ 0 & V_A + 1 \end{pmatrix}.$$

By imposing $V_A = V - 1$ we obtain after averaging over the outgoing pulses a thermal state of variance V , as shown in Fig. 7.17.

The Protocol The quantum communication consists in repeating the following steps for all the pulses sent from Alice to Bob:

1. Alice generates two random real number (a_x, a_p) from two independent Gaussian distributions of variance V_A ($V_A = V - 1$) and Bob generates a random bit b .
2. Alice sends a coherent state centered in $d = (a_x, a_p)$ to Bob.

3. Bob, depending on his random bit (b), measure either x or p .

After Bob has received all the pulses, the two partners proceed with the post-processing, which starts by applying sifting:

1. Bob discloses for each pulse the value of b (whether he measured x or p).
2. Alice keeps a_x or a_p depending on the value of b .

After the sifting follows reconciliation and privacy amplification algorithms in order to extract a secret key.

Protocol Implementation

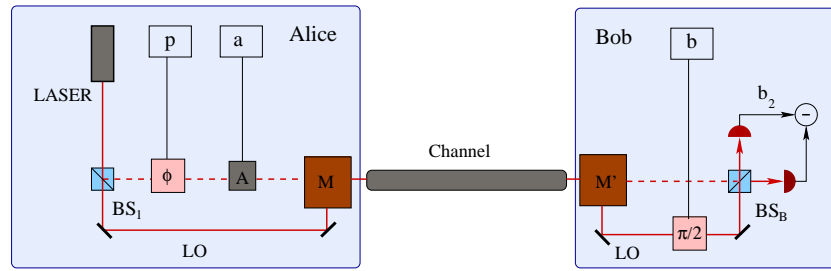


Figure 7.18: Scheme of the optical implementation of the coherent states protocol.

As shown in Fig. 7.18 the source (Alice) is based on a master laser beam that is divided at BS_1 in a classical local oscillator and a low intensity signal that is later attenuated in order to obtain a quantum coherent state with a small number of photons ($\approx 10 - 100$). The quantum coherent state is then modulated in phase (ϕ) and amplitude (A) in order to get a coherent state centered in (a_x, a_p) . Then the quantum signal and the local oscillator travel through the same fiber to Bob multiplexed (M) in time. At Bob station the two signal are first demultiplexed (M') and subsequently measured using homodyne, where Bob applies a random $\pi/2$ phase to the local oscillator depending on the quadrature he wants to measure (b).

No Basis Switching Protocol

In the previous protocol Alice generates two random real numbers but uses only one to generate the secret key. Interestingly, one can modify the coherent states protocols [96] in order to use both values, as shown in [194]. The idea is to replace Bob homodyne measurement by an heterodyne detection, where the incoming beam is divided in two using a balanced beamsplitter and we measure x on one and p on the other using homodyne detection, as shown in Fig. 7.19.

The Protocol The quantum communication step of the protocol consists in repeating the following steps for all the pulses sent from Alice to Bob, as shown in Fig. 7.19.

1. Alice generates two random real numbers (a_x, a_p) from two independent Gaussian distribution of variance V_A ($V_A = V - 1$).
2. Alice sends a coherent state centered in $d = (a_x, a_p)$ to Bob.
3. Bob applies an heterodyne measurement which extracts information on x and p .

After the quantum communication step Alice and Bob directly proceed with reconciliation and privacy amplification, as no sifting is needed during the post-processing.

Protocol Implementation

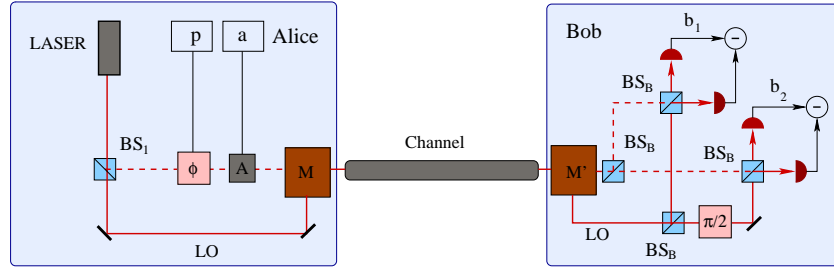


Figure 7.19: Scheme of the optical implementation of the no basis switching protocol.

As shown in Fig. 7.19 the only difference with the implementation of the coherent states protocol (Fig. 7.18) [96] is Bob measurement, where he splits the quantum signal (and the local oscillator) on two beams using a balanced beamsplitter (BS_B) and applies an homodyne measurement on each beam, as shown in Fig. 7.19. In order to measure the quadrature p Bob dephases the second local oscillator by $\theta_2 = \pi/2$.

7.7 CV-QKD Entanglement-Based Scheme

Even if most of the experimental implementations use prepare-and-measure schemes, the theoretical analysis is usually done using an entanglement-based scheme, as it simplifies the calculations. The previously presented protocols are continuous variable *prepare-and-measure* schemes, defined by a classical-quantum source

$$\rho_{aB_0} = \int da p(a) |a\rangle\langle a| \otimes |\psi_a\rangle\langle\psi_a|_{B_0}, \quad (7.30)$$

where Alice prepares according to some random number (a) a given quantum state $|\psi_a\rangle\langle\psi_a|_{B_0}$ and sends it to Bob. As pointed in Chapter 7, each prepare-and-measure protocol is strictly equivalent to a continuous-variable *entanglement-based* scheme, where Alice generates an entangled state $|\Psi\rangle_{AB_0}$, sends mode B_0 to Bob and measures mode A on the appropriate basis in order to project B_0 on the proper ensemble $|\psi_a\rangle\langle\psi_a|_{B_0}$.

Alice State Preparation

All the Gaussian CV-QKD protocols presented in the preceding section had a thermal state as average state ρ_{B_0} . A two-mode squeezed vacuum state (or EPR state) with null mean value $d = (0, 0)$ and covariance matrix

$$\gamma_{AB_0} = \begin{pmatrix} \gamma_A & \sigma_{AB_0} \\ \sigma_{AB_0} & \gamma_{B_0} \end{pmatrix} = \begin{pmatrix} V\mathbb{I} & \sqrt{V^2 - 1}\sigma_z \\ \sqrt{V^2 - 1}\sigma_z & V\mathbb{I} \end{pmatrix} \quad (7.31)$$

where σ_z reads

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (7.32)$$

being a purification of a thermal state, as explained in Chapter 2, we can construct the entanglement-based description of the family of protocols by starting from a EPR state and applying a Gaussian measurement on mode A .

Squeezed States Protocol Alice applying an homodyne measurement on mode A of an EPR state, as shown in Fig. 7.20, is equivalent to the prepare-and-measure scheme of the squeezed states protocol [34]. To prove it we use the partial measurement

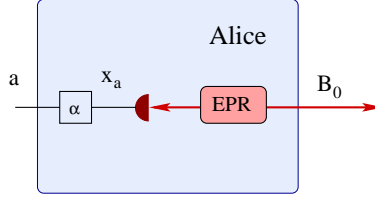


Figure 7.20: Alice applying an homodyne detection over mode A projects mode B_0 into $x(p)$ -squeezed states displaced along $x(p)$ according to a Gaussian distribution as in the prepare-and-measure scheme. Multiplies Alice's result x_a by a factor $\alpha = \sqrt{1 - 1/V^2}$ we obtain a one-to-one correspondence between the prepare-and-measure scheme and the entanglement-based scheme.

equations (2.57) and (2.58) of Section 2.3. In the case of an homodyne measurement the mean (d_{B_0}) and covariance matrix (γ_{B_0}) of mode B_0 conditioned on Alice measurement result (x_a) reads,

$$\gamma_{B_0}^{x_a} = \gamma_{B_0} - \sigma_{AB_0}^T (X \gamma_A X)^{MP} \sigma_{AB_0} = \begin{pmatrix} 1/V & 0 \\ 0 & V \end{pmatrix}, \quad (7.33)$$

and

$$d_{B_0}^{x_a} = \sigma_{AB_0}^T (X \gamma_A X)^{MP} d_A = d_{B_0} = \sqrt{1 - 1/V^2} (x_a, 0), \quad (7.34)$$

where d_A is Alice measurement result, $X = \text{diag}(1, 0, 1, 0)$ and MP denotes the pseudoinverse. This is exactly a x -squeezed vacuum state displaced over the x quadrature, in order to generate p -squeezed states Alice has just to measure the p quadrature of mode A , instead of x .

Notice that there is a one-to-one correspondence between Alice measurement result (x_a) and the mean value of mode B_0 (d_{B_0}) up to a constant $\alpha = \sqrt{1 - 1/V^2}$ ($\alpha \approx 1$ when $V \gg 1$). The variance of d_{B_0} reads,

$$\langle \Delta^2 d_{B_0} \rangle = \alpha \langle x_a^2 \rangle = V - 1/V, \quad (7.35)$$

as $\langle x_a^2 \rangle = V$. Notice that $\langle \Delta^2 d_{B_0} \rangle$ is equal to V_A , the variance of Alice modulation (see Section 7.6). By adding a classical multiplication stage $a = \alpha x_a$ after Alice measurement we observe a one-to-one correspondence between the prepare-and-measure scheme and the entanglement-based scheme.

General Preparation

Following [95], a more general measurement consists in Alice applying a generalized heterodyne detection, where we use an unbalanced beamsplitter of transmittance T_A , as shown in Fig. 7.21. Before the beamsplitter of transmittance T_A the tripartite state ACB_0 reads,

$$\gamma_{ACB_0} = \begin{pmatrix} V\mathbb{I} & 0 & \sqrt{V^2 - 1}\sigma_z \\ 0 & \mathbb{I} & 0 \\ \sqrt{V^2 - 1}\sigma_z & 0 & V\mathbb{I} \end{pmatrix}, \quad (7.36)$$

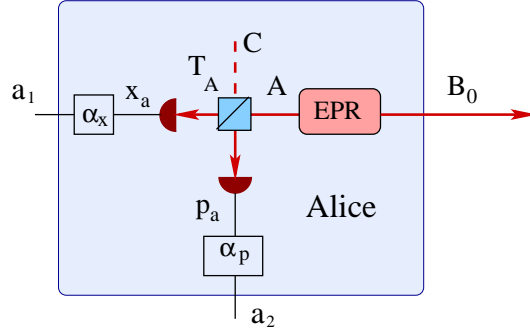


Figure 7.21: Alice applying a generalized heterodyne detection (T_A) over mode A projects mode B_0 on general squeezed states displaced among a bi-variate Gaussian distribution. By adding a classical multiplication stage $a = \alpha(x_a, p_a)$, after Alice measurement we observe a one-to-one correspondence between the prepare-and-measure scheme and the entanglement-based scheme.

where C is the second input of the beamsplitter which is initially in the vacuum. The beamsplitter transforms the tripartite state as

$$\gamma'_{ACB_0} = [S_{AC}^{BS} \otimes \mathbb{I}_{B_0}]^T \gamma_{ACB_0} [S_{AC}^{BS} \otimes \mathbb{I}_{B_0}] \quad (7.37)$$

where S_{AC}^{BS} is the symplectic transformation of the beamsplitter (T_A),

$$S_{AC}^{BS} = \begin{pmatrix} \sqrt{T_A} \mathbb{I} & \sqrt{1-T_A} \mathbb{I} \\ -\sqrt{1-T_A} \mathbb{I} & \sqrt{T_A} \mathbb{I} \end{pmatrix}. \quad (7.38)$$

Alice measurement on modes A (x_a) and C (p_a) projects mode B_0 (using equation (7.33)) into a Gaussian state of covariance matrix,

$$\gamma_{B_0}^{(x_a, p_a)} = \begin{pmatrix} \frac{\mu V + 1}{V + \mu} & 0 \\ 0 & \left(\frac{\mu V + 1}{V + \mu} \right)^{-1} \end{pmatrix}, \quad (7.39)$$

where

$$\mu = \frac{1 - T_A}{T_A}, \quad (7.40)$$

and mean value (using equation (7.34)),

$$d_{B_0}^{(x_a, p_a)} = \left(\frac{\sqrt{T_A(V^2 - 1)}}{T_A V + (1 - T_A)} x_a, \frac{\sqrt{(1 - T_A)(V^2 - 1)}}{(1 - T_A)V + T_A} p_a \right), \quad (7.41)$$

which is a squeezed state displaced along a bi-variate Gaussian distribution. For $T_A = 1$ we recover the entangled-based description of the squeezed states protocol.

Coherent states protocol If we fix $T_A = 1/2$ the covariance matrix and mean of the projected state of mode B_0 read,

$$\gamma_{B_0}^{(x_a, p_a)} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{and} \quad d_{B_0}^{(x_a, p_a)} = \sqrt{2 \frac{V-1}{V+1}} (x_a, p_a), \quad (7.42)$$

which is a coherent state centered at $d_{B_0}^{(x_a, p_a)}$. We observe that the variance of $d_{B_0}^{(x_a, p_a)}$, using $\langle x_a^2 \rangle = (V+1)/2$ reads,

$$\langle \Delta^2 d_{B_0} \rangle = V - 1 = V_A, \quad (7.43)$$

being exactly the variance of the bi-variate Gaussian modulation of protocol [96]. Similarly as for the squeezed states protocol, adding a classical multiplication stage $a = \alpha(x_a, p_a)$, as shown in Fig. 7.21, after Alice measurement we observe a one-to-one correspondence between the "prepare-and-measure" scheme and the entanglement-based scheme.

Entanglement-Based Scheme of CV-QKD

One can unify the description of all the previously proposed protocols into a single entanglement-based scheme, as shown in Fig. 7.22, which simplifies the theoretical calculations. Depending on Alice measurement the prepared states are:

- Squeezed states, if Alice applies a homodyne measurement ($T_A = 1$)
- Coherent states, if Alice applies a heterodyne measurement ($T_A = 1/2$)

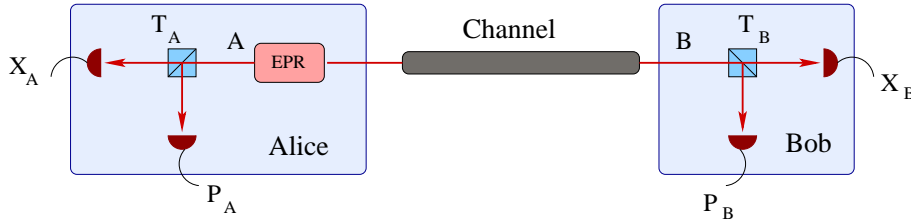


Figure 7.22: Entanglement-based scheme that generalizes all the previously presented protocols. Depending on $T_A = \{1, 1/2\}$ Alice generates squeezed state or coherent states and depending on $T_B = \{1, 1/2\}$ Bob applies homodyne or heterodyne detection.

Bob has the choice between applying two different measurements:

- Homodyne measurement on mode B ($T_A = 1$)
- Heterodyne measurement on mode B ($T_A = 1/2$)

We observe that there are four different optical implementations depending on the values of $T_A, T_B = \{1, 1/2\}$, which combined to two different classical post-processing (Direct and Reverse Reconciliation) generate eight possible protocols. Remark that there is one combination that has not been proposed yet in the literature, the implementation where Alice sends squeezed states ($T_A = 1$) and Bob applies heterodyne detection ($T_B = 1/2$). The probable reason for the lack of interest on this protocol is that it is a noisy version of the protocol based on squeezed states ($T_A = 1$) and homodyne detection ($T_B = 1$). Surprisingly, this protocol can be very interesting in some situations as we show in Chapter 9.

Gaussian Channel

In Chapters 8 and 9 we will show that the optimal Individual and Collective attacks are always a Gaussian CP map. It is then enough to consider the bipartite state ρ_{AB} shared by Alice and Bob in the entanglement-based description as being Gaussian, which simplifies very much the calculations.

In this dissertation the security analysis of the protocol would be done assuming that Eve does a *passive* attack, where she replaces the real connection between Alice and Bob by a unitary operation that mimics the channel between Alice and Bob when we trace Eve's modes. An optical fiber can be modeled by a thermal noise channel of

transmittance T and noise referred to the input χ . Then the covariance matrix of the state ρ_{AB} reads,

$$\gamma_{AB} = \begin{pmatrix} V\mathbb{I} & \sqrt{T(V^2-1)}\sigma_z \\ \sqrt{T(V^2-1)}\sigma_z & V(T+\chi)\mathbb{I} \end{pmatrix}. \quad (7.44)$$

But in QKD we have to assume that Eve can apply any attack, an *active* Eve could certainly decide to apply a different attack generating a different γ_{AB} than that of equation (7.44). The calculation for general γ_{AB} in the case of collective attacks is numerically as simple as those of the thermal noise channel, but has much more than two free parameters (T and χ) making the comparison of the protocols more difficult. For the sake of simplicity, in this dissertation we will give the results only for thermal noise channels as it gives us an intuition on the efficiency of the protocols, but we stress that its generalization to more general channels is easy.

Chapter 8

CV-QKD: Individual Attacks

8.1 Introduction

In this chapter we are going to study the security of Gaussian Quantum Key Distribution protocols presented in the previous chapter against individual attacks, where Eve information is quantified by the accessible information defined in Chapter 6. In a QKD scheme we consider that Eve (E) holds the purification of Alice and Bob quantum system ρ_{AB} . The pure tripartite state shared by Alice, Bob and Eve reads $|\Psi\rangle_{ABE}$, which is completely defined by the eigenvalues of ρ_{AB} . This implies that any function on the system ABE , such as Eve's accessible information on Alice measurement output a , is in fact a function of ρ_{AB} . We will use the notation $I_{acc}^E(\rho_{AB}; M_A)$ for Eve's accessible information on Alice data (a resulting from the POVM M_A),

$$I_{acc}^E(\rho_{AB}; M_A) = \max_{M_E} I(\rho_{AE}; M_A, M_E), \quad (8.1)$$

to stress that it only depends on ρ_{AB} , where we use the upperscript E on I^E to differentiate it from Alice and Bob accessible information $I_{acc}(\rho_{AB}; M_A)$.

Secret Key Rates

In the scenario of individual attack the achievable Direct Reconciliation (DR) secret key rate reads,

$$K^{DR}(\rho_{AB}) = I(\rho_{AB}; M_A, M_B) - I_{acc}^E(\rho_{AB}; M_A). \quad (8.2)$$

The Reverse Reconciliation (RR) achievable secret key rate reads,

$$K^{RR}(\rho_{AB}) = I(\rho_{AB}; M_A, M_B) - I_{acc}^E(\rho_{AB}; M_B). \quad (8.3)$$

Entropy of a Measurement Result

In some situations we will use the notation

$$H(\rho_X; M_X), \quad (8.4)$$

to express the entropy of the measurement result x after applying the POVM M_X to the quantum state ρ_X .

8.2 Optimality of Gaussian Individual Attacks

To prove the optimality of Gaussian individual attacks, we need first to remind a very useful theorem, recently proven in [198]. Subsequently we will prove the optimality of Gaussian attacks against Direct Reconciliation Gaussian protocols. Its extension to Reverse Reconciliation being straightforward, one simply needs to interchange the roles of Alice and Bob.

Optimality of Gaussian States

Let us first sketch the proof in [198] for bipartite states ρ_{AB} with null mean value. Let f be a function satisfying the properties

1. continuity in trace norm: if $\|\rho_{AB}^{(n)} - \rho_{AB}\|_1 \rightarrow 0$ when $n \rightarrow \infty$, then $f(\rho_{AB}^{(n)}) \rightarrow f(\rho_{AB})$,
2. invariance under local ‘‘Gaussification’’ unitaries (defined below):
 $f(U_G^\dagger \otimes U_G^\dagger \rho_{AB}^{\otimes N} U_G \otimes U_G) = f(\rho_{AB}^{\otimes N})$,
3. strong sub-additivity: $f(\rho_{A_1 \dots N B_1 \dots N}) \leq f(\rho_{A_1 B_1}) + \dots + f(\rho_{A_N B_N})$ with equality if $\rho_{A_1 \dots N B_1 \dots N} = \rho_{A_1 B_1} \otimes \dots \otimes \rho_{A_N B_N}$.

Then, for every bipartite state ρ_{AB} with covariance matrix γ_{AB} , we have that

$$f(\rho_{AB}) \leq f(\rho_{AB}^G), \quad (8.5)$$

where ρ_{AB}^G is the Gaussian state with the same γ_{AB} . The proof can be summarized by

$$\begin{aligned} f(\rho_{AB}) &\stackrel{3}{=} \frac{1}{N} f(\rho_{AB}^{\otimes N}) \stackrel{2}{=} \frac{1}{N} f(\tilde{\rho}_{A_1 \dots N B_1 \dots N}) \\ &\stackrel{3}{\leq} \frac{1}{N} \sum_{k=1}^N f(\tilde{\rho}_{A_k B_k}) \stackrel{1, \star}{\simeq} f(\rho_{AB}^G), \end{aligned} \quad (8.6)$$

where the superscripts label the assumptions used in each step, while

$$\tilde{\rho}_{A_1 \dots N B_1 \dots N} \equiv U_G^\dagger \otimes U_G^\dagger \rho_{AB}^{\otimes N} U_G \otimes U_G. \quad (8.7)$$

The \star stands for the use of a central limit result for quantum states (see [198] for details). The Gaussification unitary U_G is a local passive operation that applies the

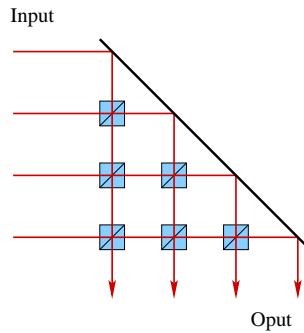


Figure 8.1: The Gaussification U_G can be realized with the following network of beam splitters and phase shifters.

same transformation to both quadratures independently:

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}_{out} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{\otimes m} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_N \end{bmatrix}_{in}, \quad (8.8)$$

where $N = 2^m$. The Gaussification U_G can be realized with a network of beam splitters and phase shifters, as shown in Fig.. Importantly for what follows, the x and p quadratures of all N modes are thus not mixed via Gaussification.

Optimality of Gaussian Individual Attacks

The core of our proof now consists in combining this extremality result with a generalized version of the entanglement-based version of CV-QKD, where Alice and Bob apply a general Gaussian measurement which does not mix both quadrature, as shown in Fig. 8.2, supplemented with the physical model of measurement. In a real implementation

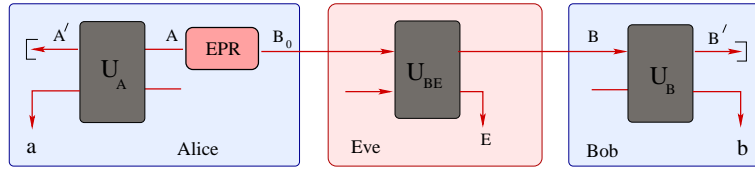


Figure 8.2: Entanglement-based scheme for CV-QKD. Alice's preparation is modelled by a measurement U_A on her half of an EPR pair. The channel is modelled by an unitary interaction between mode B and Eve's ancillae E . Finally, Bob's measurement is modelled by U_B . The only constraint on the measurement is that both unitary interactions U_A and U_B should not mix both quadratures.

of CV-QKD protocol Alice and Bob mutual information $S(a:b)$ is fixed by the data obtained by the two partners and the efficiency of the reconciliation. Then, in order to prove the optimality of Gaussian attacks it is enough to prove that Eve information $I_{acc}^E(\rho_{AB}; M_A)$ is maximized when she applies a Gaussian map.

Proof

We are going to prove that $I_{acc}^E(\rho_{AB}; M_A)$ satisfies the three conditions of the Gaussian extremality theorem. For this, we need to define the extension of this function over $2N$ modes ($\mathbf{A} = A_1, A_2, \dots, A_N$, $\mathbf{B} = B_1, B_2, \dots, B_N$), namely

$$I_{acc}^E(\rho_{\mathbf{AB}}; M_A^{\otimes N}) = \max_{M_E} I(\rho_{\mathbf{AE}}; M_A^{\otimes N}, M_E), \quad (8.9)$$

where Alice applies the same measurement M_A on each mode, and Eve applies the optimal POVM M_E on ρ_E which is the purification of $\rho_{\mathbf{AB}}$. Note that Eq. (8.9) restricts to Eq. (8.1) when $N = 1$.

Continuity If $\|\rho_{\mathbf{AB}}^{(n)} - \rho_{\mathbf{AB}}\|_1 \leq \epsilon$, using Uhlmann's theorem and well-known relations between the fidelity and trace distance (see Appendix I), we can find a purification $|\Psi\rangle_{\mathbf{ABE}}^{(n)}$ ($|\Psi\rangle_{\mathbf{ABE}}$) of $\rho_{\mathbf{AB}}^{(n)}$ ($\rho_{\mathbf{AB}}$) such that

$$\|\hat{\Psi}_{\mathbf{ABE}}^{(n)} - \hat{\Psi}_{\mathbf{ABE}}\|_1 \leq 2\sqrt{\epsilon}. \quad (8.10)$$

We define the difference of accessible informations over $\rho_{\mathbf{AB}}^{(n)}$ and $\rho_{\mathbf{AB}}$ as,

$$\Delta I_{acc}^E = I_{acc}^E(\rho_{\mathbf{AB}}^{(n)}; M_A^{\otimes N}) - I_{acc}^E(\rho_{\mathbf{AB}}; M_A^{\otimes N}). \quad (8.11)$$

This is equivalent by equation (8.9) to

$$\Delta I_{acc}^E = \max_{M_E} I(\rho_{\mathbf{AE}}^{(n)}; M_A^{\otimes N}, M_E) - \max_{M_E} I(\rho_{\mathbf{AE}}; M_A^{\otimes N}, M_E). \quad (8.12)$$

Eve's optimal POVM over $\rho_{\mathbf{AE}}^{(n)}$ (M_E^{opt}) is not necessarily optimal on $\rho_{\mathbf{AE}}$, we obtain then,

$$\Delta I_{acc}^E \leq I(\rho_{\mathbf{AE}}^{(n)}; M_A^{\otimes N}, M_E^{opt}) - I(\rho_{\mathbf{AE}}; M_A^{\otimes N}, M_E^{opt}) \quad (8.13)$$

$$\leq \left[H(\rho_{\mathbf{A}}^{(n)}; M_A^{\otimes N}) - H(\rho_{\mathbf{A}}; M_A^{\otimes N}) \right] \quad (8.14)$$

$$+ \left[H(\rho_{\mathbf{E}}^{(n)}; M_E^{opt}) - H(\rho_{\mathbf{E}}; M_E^{opt}) \right] \quad (8.15)$$

$$- \left[H(\rho_{\mathbf{AE}}^{(n)}; M_A^{\otimes N}, M_E^{opt}) - H(\rho_{\mathbf{AE}}; M_A^{\otimes N}, M_E^{opt}) \right] \quad (8.16)$$

Then, considering that partial trace can only decrease the trace norm [136], Eq. (8.10) combined with our physical model of measurement and the continuity of von Neumann entropies together with (8.16) implies $\Delta I_{acc}^E \leq \delta$. Using the same argument but with the POVM optimal for $\rho_{\mathbf{AB}}$ we obtain $-\Delta I_{acc}^E \leq \delta$, giving $|\Delta I_{acc}^E| \leq \delta$, implying the continuity of I_{acc}^E . \square

Invariance under Local Gaussification Unitaries Applying the local Gaussification operation $U_G \otimes U_G$ on the product states $|\psi\rangle_{ABE}^{\otimes N}$ (as shown in Fig. 8.3 for $N = 2$), we obtain the state $|\tilde{\psi}\rangle_{\mathbf{ABE}}$. After the measurements on Alice's and Bob's sides, the

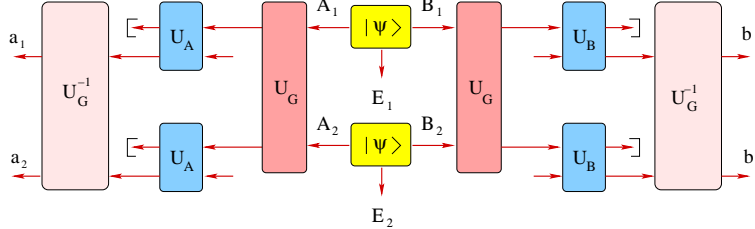


Figure 8.3: Invariance under local “Gaussification” unitaries: U_G can be interchanged with the measurement U_A , then U_G^{-1} and U_G cancel each other.

state becomes $\tilde{\rho}_{\mathbf{abE}}$. But because the Gaussian measurement (usually homodyne or heterodyne) does not mix x and p quadratures, the measurement and the Gaussification operation can be interchanged by applying $U_G^\dagger \otimes U_G^\dagger$ on modes \mathbf{a} and \mathbf{b} we recover the state $\rho_{\mathbf{abE}}^{\otimes N}$, which coincides with the state obtained by directly measuring $|\psi\rangle_{ABE}^{\otimes N}$ without Gaussification. Since the two states $\tilde{\rho}_{\mathbf{ab}}$ and $\rho_{\mathbf{ab}}^{\otimes N}$ are related by a local unitary operation $U_G^\dagger \otimes U_G^\dagger$ and since the von Neumann entropies appearing in $I_{acc}^E(\rho_{\mathbf{AB}}; M_a^{\otimes N})$ are invariant under (any) local unitaries, we obtain the invariance of $I_{acc}^E(\rho_{\mathbf{AB}}; M_a^{\otimes N})$ under local Gaussification unitaries. \square

Strong Subadditivity We will restrict the proof to two modes on each side, $A_{1,2}$ and $B_{1,2}$, as shown in Fig. 8.4, where E is the purification of $A_{1,2}B_{1,2}$. The generalization to $N > 2$ being straightforward.

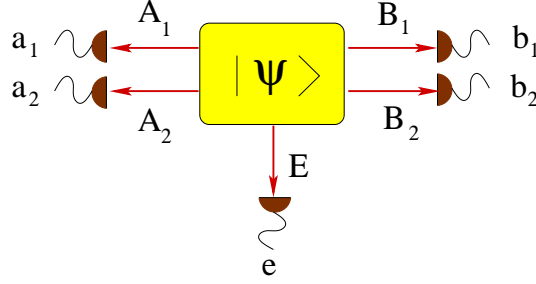


Figure 8.4: Alice and Bob share a quantum state $A_{1,2}B_{1,2}$ where Alice applies independent measurements on each mode $M_{A_1} \otimes M_{A_2}$, and so does Bob $M_{B_1} \otimes M_{B_2}$. Eve holds the purification (E) of $A_{1,2}B_{1,2}$. Remark that the purification of A_1B_1 (A_2B_2) is A_2B_2E (A_1B_1E).

Before continuing, let's stress a property of Shannon mutual entropy

$$H(a_1a_2 : e) = H(a_1a_2) - H(a_1a_2|e), \quad (8.17)$$

which can be rewritten,

$$\begin{aligned} H(a_1a_2 : e) &= \underbrace{H(a_1a_2)}_{\leq H(a_1)+H(a_2)} - [\underbrace{H(a_1|a_2e) + H(a_2|a_1e)}_{\geq H(a_1|a_2b_2e)+H(a_2|a_1b_1e)} + \underbrace{H(a_1 : a_2|e)}_{\geq 0}] \\ &\leq H(a_1) + H(a_2) - H(a_1|a_2b_2e) - H(a_2|a_1b_1e). \end{aligned} \quad (8.18)$$

as a consequence of subadditivity, strong subadditivity of entropy and using the fact that conditioning can only decrease the conditional entropy. This bound can be rewritten as,

$$H(a_1a_2 : e) \leq H(a_1 : a_2b_2e) + H(a_2 : a_1b_1e), \quad (8.19)$$

which using the notation used for the accessible information reads,

$$\begin{aligned} I(\rho_{A_1A_2E}; M_{A_1} \otimes M_{A_2}, M_E) &\leq I(\rho_{A_1E_1}; M_{A_1}, M_{A_2} \otimes M_{B_2} \otimes M_E) \\ &\quad + I(\rho_{A_2E_2}; M_{A_2}, M_{A_1} \otimes M_{B_1} \otimes M_E). \end{aligned} \quad (8.20)$$

Considering Eve's optimal measurement over system E that reaches the accessible information over (a_1, a_2) resulting from Alice measurement ($M_{A_1} \otimes M_{A_2}$) on system A_1A_2 , and knowing that the purification of A_1B_1 (A_2B_2) is $E_1 = A_2B_2E$ ($E_2 = A_1B_1E$) we can rewrite (8.19) as

$$\begin{aligned} I_{acc}(\rho_{A_1A_2E}; M_{A_1} \otimes M_{A_2}) &\leq I(\rho_{A_1E_1}; M_{A_1}, M_{A_2} \otimes M_{B_2} \otimes M_E^{opt}) \\ &\quad + I(\rho_{A_2E_2}; M_{A_2}, M_{A_1} \otimes M_{B_1} \otimes M_E^{opt}). \end{aligned} \quad (8.21)$$

Considering that the POVM $M_{A_2} \otimes M_{B_2} \otimes M_E^{opt}$ ($M_{A_1} \otimes M_{B_1} \otimes M_E^{opt}$) is not necessarily the measurement optimizing the classical correlation between a_1 (a_2) and A_2B_2E (A_1B_1E) we obtain,

$$I_{acc}(\rho_{A_1A_2E}; M_{A_1} \otimes M_{A_2}) \leq I_{acc}(\rho_{A_1E_1}; M_{A_1}) + I_{acc}(\rho_{A_2E_2}; M_{A_2}). \quad (8.22)$$

which is exactly the strong subadditivity of Eve's accessible information. \square

Thus, using Eq. (8.5), we have proved that for all bipartite quantum states ρ_{AB} with covariance matrix γ_{AB} , one has $I_{acc}^E(\rho_{AB}; M_A) \leq I_{acc}^E(\rho_{AB}^G; M_A)$. This means that $I_{acc}^E(\rho_{AB}^G; M_A)$ is an upper bound on Eve accessible information on Alice data for

any protocol (even non-Gaussian) and individual attack (including non-Gaussian). The only requirement for this result to hold is that Alice and Bob measurements commute with the Gaussification operation U_G and that Alice and Bob use the second-order moments of the quadratures in order to calculate this bound. In particular, for the Gaussian-modulation protocols of [34, 96, 94, 194] presented in the previous chapter, Eve's optimal attack is a Gaussian attack, in which case the bound is saturated. Note that the proof concerns DR, but its extension to RR is straightforward: one simply needs to interchange Alice and Bob roles.

8.3 Security Analysis using Uncertainty Relations

The calculation of the accessible information is not trivial. In order to calculate it for the eight Gaussian protocols presented in the previous chapter we will proceed in two steps; Firstly, we will use continuous variable entropic uncertainty relations to upperbound Eve's information; Secondly, we will try to find the implementation that saturates the bound.

Since Gaussian attacks are optimal, we consider in what follows that Eve affects a Gaussian channel. Consequently, the quantum state ρ_{AB} before Alice and Bob's measurements can be assumed to be a Gaussian two-mode state with a zero mean value and a covariance matrix γ_{AB} . Usual Gaussian channels, such as optical fibers, add a symmetric and uncorrelated noise in both quadratures x and p (including, of course, the loss-induced noise), so that we will only consider symmetric channels without x - p correlations in what follows. Since the EPR state (two-mode squeezed state) is also symmetric and exhibits no correlations between x and p , we can write the resulting covariance matrix in a block-diagonal form as

$$\gamma_{AB} = \begin{pmatrix} \gamma_{AB}^x & 0 \\ 0 & \gamma_{AB}^p \end{pmatrix}, \quad (8.23)$$

with

$$\gamma_{AB}^{x(p)} = \begin{pmatrix} V & \pm \sqrt{T(V^2 - 1)} \\ \pm \sqrt{T(V^2 - 1)} & T(V + \chi) \end{pmatrix} \quad (8.24)$$

where the signs $+$ and $-$ correspond to γ_{AB}^x and γ_{AB}^p , respectively. Here, V is the variance of Alice's output thermal state, while T and $\chi = (1 - T)/T + \epsilon$ are the transmittance and noise referred to the input of the Gaussian channel [the term $(1 - T)/T$ stands for the loss-induced vacuum noise, while ϵ is the excess noise referred to the input].

Entropic Uncertainty Relations

In order to address the security of the protocols, we may, without loss of generality, assume that Eve holds the purification of the quantum state ρ_{AB} . By measuring their systems, Bob and Eve then project Alice's share of the joint pure state $|\Psi_{ABE}\rangle$ onto another pure state¹. Applying the Heisenberg uncertainty relation on the pure state held by Alice (conditioning on Bob and Eve's measurements), we have

$$V_{X_A|E} V_{P_A|B} \geq 1, \quad (8.25)$$

where X_A and P_A are the canonically conjugate quadratures of Alice's mode and $V_{X|Y}$ is the conditional variance measuring the remaining uncertainty on X after the

¹We may indeed always assume that Eve performs a measurement based on a *rank-one* Positive Operator Valued Measure (POVM), so that the resulting state is pure. Otherwise, she would just need to disregard a part of her measuring system.

measurement of Y ,

$$V_{X|Y} = \langle x^2 \rangle - \frac{\langle xy \rangle^2}{\langle y^2 \rangle}, \quad (8.26)$$

expressed in shot-noise units. Equation (8.25) also has a symmetric counterpart that reads,

$$V_{P_A|E} V_{X_A|B} \geq 1. \quad (8.27)$$

Since we focus on a symmetric noise in x and p , Eqs. (8.25) and (8.27) can be unified into a single uncertainty relation

$$V_{A|E} V_{A|B} \geq 1, \quad (8.28)$$

where A stands for any quadrature (X_A or P_A) of Alice's mode. This inequality will be used to put a lower bound on the uncertainty of Eve's estimate of the key in Direct Reconciliation (DR). Similarly, in Reverse Reconciliation (RR), one can derive a dual inequality

$$V_{B|E} V_{B|A} \geq 1, \quad (8.29)$$

where B stands for any quadrature of Bob's mode. This will be used to put a lower bound on the uncertainty of Eve's estimate of the key in RR.

Now, we will derive lower bounds on the secret key rates using the above uncertainty relations on the variances.

Squeezed States and Homodyne Detection

The protocol based on squeezed states and homodyne measurement [44] is equivalent to an entanglement based scheme where Alice and Bob apply homodyne measurements over modes A and B respectively, as shown in Fig. 8.5. Restricting to individual

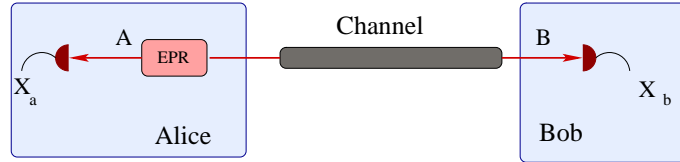


Figure 8.5: Entanglement based scheme of the protocol based on Alice sending squeezed states and Bob applying homodyne detection. Alice generation of squeezed states is replaced by an EPR state where Alice applies an homodyne detection on one half of the EPR and the other half is sent to Bob.

attacks and one-way reconciliation, the DR and RR secret key rates read

$$K_{\text{DR}} = H(A|E) - H(A|B), \quad (8.30)$$

$$K_{\text{RR}} = H(B|E) - H(B|A), \quad (8.31)$$

where E stands for Eve's optimal measurement maximizing her information (which is not necessarily the same in DR and RR). If we assume that the channel is Gaussian, we can express the conditional entropies in Eqs. (8.30) and (8.31) in terms of conditional variances,

$$H(X|Y) = \frac{1}{2} \log V_{X|Y}, \quad (8.32)$$

where the log is to the base 2 and entropy is expressed in bits. The above Heisenberg inequalities on conditional variances directly translate into bounds on the secret key rates.

Direct Reconciliation

Using equations (8.26) and (8.23) one obtains,

$$V_{A|B} = \frac{V\chi + 1}{V + \chi}, \quad (8.33)$$

which together with equation (8.30) and the Heisenberg uncertainty relation (8.28) gives

$$K_{\text{DR}} = \frac{1}{2} \log \left[\frac{V_{A|E}}{V_{A|B}} \right] \geq \log \left[\frac{1}{V_{A|B}} \right] = \log \left[\frac{V + \chi}{V\chi + 1} \right], \quad (8.34)$$

as shown in [34].

Reverse Reconciliation

Using equations (8.26) and (8.23) one obtains,

$$V_{B|A} = T(\chi + 1/V), \quad (8.35)$$

which together with equation (8.31) and the Heisenberg uncertainty relation (8.29) gives

$$K_{\text{RR}} = \frac{1}{2} \log \left[\frac{V_{B|E}}{V_{B|A}} \right] \geq \log \left[\frac{1}{V_{B|A}} \right] = \log \left[\frac{1}{T(\chi + 1/V)} \right], \quad (8.36)$$

as shown in [95].

Implementation of the Optimal Attack

In order to saturate (8.34, 8.36) Eve can use the so called *entangling cloner* [95]. In an

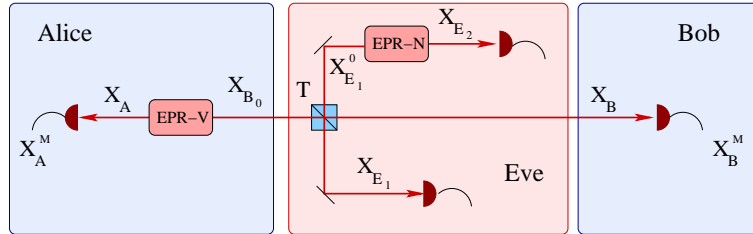


Figure 8.6: Eve replaces the usual Gaussian channel of transmittance T and excess noise referred to the input χ by half of an EPR pair (N) that she mixes with Alice signal into a beamsplitter of transmittance T . The value of N is tuned in order to inject the same noise as in the original channel $\chi = (1 - T)N/T$.

entangling cloner attack Eve replaces the Gaussian channel with transmittance T and excess noise referred to the input χ by an EPR pair of variance N and a beamsplitter of transmittance T . Half of the EPR is mixed with the state sent by Alice \hat{x}_{B_0} in the beamsplitter. Alice and Bob having only access to half of the EPR, they see a thermal noise of variance N , where N is tuned to match the noise of the real channel. The other half of the EPR will be used by Eve to reduce her uncertainty on the noise added by the channel. The channel being symmetric for x and p quadratures, we restrict our study to the x quadrature. The quantum signal received by Bob reads,

$$\hat{x}_B = \sqrt{T}\hat{x}_{B_0} + \sqrt{1-T}\hat{x}_{E_1^0}, \quad (8.37)$$

where using (8.23) the variance reads,

$$\langle \hat{x}_B^2 \rangle = T \left[V + \frac{1-T}{T} N \right], \quad (8.38)$$

as $\langle \hat{x}_{B_0}^2 \rangle = V$ and $\langle \hat{x}_{B_0} \hat{x}_{E_1} \rangle = 0$. In order to have the same variance $\langle \hat{x}_B^2 \rangle$ as in the real channel (8.23) Eve has to fix N in order to satisfy,

$$\chi = \frac{1-T}{T} N. \quad (8.39)$$

Eve stores her two ancillary systems, E_1 and E_2 , in two quantum memories. After Alice and Bob reveal the selected basis during the sifting step, Eve measures the right quadrature on systems E_1 and E_2 . The measurement on E_2 will allow her to decrease the noise added by mode E_1 .

Direct Reconciliation The quadrature of mode E_1 , his variance and correlations with A read,

$$\hat{x}_{E_1} = \sqrt{T} \hat{x}_{E_1}^0 - \sqrt{1-T} \hat{x}_{B_0}, \quad (8.40)$$

$$\langle \hat{x}_{E_1}^2 \rangle = TV_{E_1} + (1-T)V, \quad (8.41)$$

$$\langle \hat{x}_A \hat{x}_{E_1} \rangle^2 = (1-T)(V^2 - 1). \quad (8.42)$$

If Eve uses only her measurement on mode E_1 to estimate \hat{x}_A her uncertainty, using Eq. 8.26 reads,

$$V_{A|E_1} = \frac{V + \left[\frac{1-T}{TV_{E_1}} \right]}{V \left[\frac{1-T}{TV_{E_1}} \right] + 1}, \quad (8.43)$$

where $V_{E_1} = N$. But Eve's measurement of \hat{x}_{E_2} on the EPR allows her to reduce her uncertainty on the other half of the EPR (\hat{x}_{E_1}), obtaining

$$V_{E_1|E_2} = \frac{1}{N}. \quad (8.44)$$

After replacing V_{E_1} in Eq. (8.43) by Eq. (8.44), Eve uncertainty on \hat{x}_A reads,

$$V_{A|E_1, E_2} = \frac{V + \left[\frac{(1-T)N}{T} \right]}{V \left[\frac{(1-T)N}{T} \right] + 1} = \frac{V + \chi}{V\chi + 1}, \quad (8.45)$$

which is exactly the $V_{A|E}$ saturating the bound (8.34), which shows that the entangling cloner is an optimal attack against the protocol based on squeezed states and homodyne detection and Direct Reconciliation.

Reverse Reconciliation The correlations between E_1 and B read,

$$\langle \hat{x}_B \hat{x}_{E_1} \rangle^2 = \sqrt{T(1-T)}(V_{E_1} - V). \quad (8.46)$$

If Eve only uses her measurement on mode E_1 to estimate \hat{x}_B her uncertainty reads (using Eq. 8.26),

$$V_{B|E_1} = \frac{1}{T \left[\left(\frac{1-T}{TV_{E_1}} \right) + 1/V \right]}, \quad (8.47)$$

where $V_{E_1} = N$. But Eve's measurement of \hat{x}_{E_2} allows her to reduce her uncertainty on the other half of the EPR ($V_{E_1|E_2} = 1/N$), obtaining

$$V_{B|E_1, E_2} = \frac{1}{T[(\frac{(1-T)N}{T}) + 1/V]} = \frac{1}{T[\chi + 1/V]}, \quad (8.48)$$

which is exactly the $V_{B|E}$ saturating the bound (8.36), which shows that the entangling cloner is also an optimal attack against the Reverse Reconciliation version of the protocol based on squeezed states and homodyne detection.

Coherent States and Homodyne Detection

The protocol based on coherent states and homodyne measurement [96] is equivalent to an entanglement based scheme where Alice applies an heterodyne measurement over mode A and Bob applies homodyne measurements over modes B , as shown in Fig. 8.7.

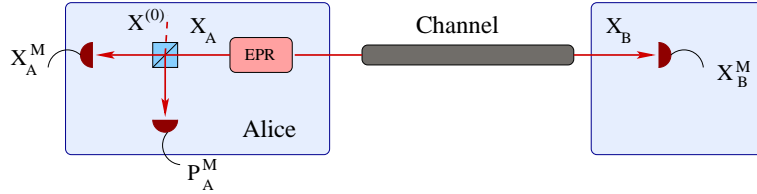


Figure 8.7: Entanglement based scheme of the protocol based on Alice sending coherent states and Bob applying homodyne detection. Alice generation of coherent states is replaced by an EPR state where Alice applies an heterodyne detection on one half of the EPR and the other half is sent to Bob.

Restricting to individual attacks and one-way reconciliation, the DR and RR secret key rates read

$$K_{\text{DR}} = H(A^M|E) - H(A^M|B), \quad (8.49)$$

$$K_{\text{RR}} = H(B|E) - H(B|A^M). \quad (8.50)$$

Note that we use the variable A^M here (not A), since in this protocol Alice does not measure one single quadrature but a pair of conjugate quadratures [A^M stands for the measurement of one quadrature of mode A , given that the conjugate quadrature is simultaneously measured]. The quadrature measured by the heterodyne detection reads,

$$\hat{x}_{A^M} = \frac{1}{\sqrt{2}}[\hat{x}_A + \hat{x}^{(0)}], \quad (8.51)$$

where $\hat{x}^{(0)}$ is the x quadrature of the vacuum noise added during the heterodyne measurement (see Fig. 8.7). One can show using equation (8.26) that the variances conditioned on an arbitrary variable Y before and after the heterodyne detection are related by,

$$V_{X_{A^M}|Y} = \frac{1}{2}[V_{X_A|Y} + 1], \quad (8.52)$$

as any quadrature Y is independent from the vacuum noise added during the measurement $\hat{x}^{(0)}$ ($\langle \hat{X}^{(0)}Y \rangle = 0$).

Direct Reconciliation

Using equations (8.52) and (8.33) for symmetric channels one obtains,

$$V_{A^M|B} = \frac{1}{2} [V_{A|B} + 1]. \quad (8.53)$$

Eve modes and the vacuum noise at the measurement being not correlated, one can then bound Eve uncertainty using equations (8.52) and (8.45),

$$V_{A^M|E} = \frac{1}{2} [V_{A|E} + 1] \geq \frac{1}{2} \left[\frac{1}{V_{A|B}} + 1 \right]. \quad (8.54)$$

Finally the secret key rate reads,

$$K_{\text{DR}} = \frac{1}{2} \log \left[\frac{V_{A^M|E}}{V_{A^M|B}} \right] \geq \frac{1}{2} \log \left[\frac{1}{V_{A|B}} \right] = \frac{1}{2} \log \left[\frac{V + \chi}{V\chi + 1} \right]. \quad (8.55)$$

Remark that the secret key rate obtained for Direct Reconciliation is just half of key rate for the protocol using squeezed states and homodyne measurement (8.34). This is not in contradiction with the claim in [96], as in this paper they decrease the key rate by a factor 1/2 in order to consider the loss of half of the data during the sifting procedure, due to Bob's wrong basis choices. This loss can be avoided if one considers an unbalanced choice of quadratures and waits for longer times as explained in Sec. 7.3. Then, one can asymptotically reach a sifting prefactor equal to one.

Reverse Reconciliation

In the case of Reverse Reconciliation the heterodyne measurement only affects Alice and Bob mutual information as Eve information on Bob data is not affected by the measurement at Alice's side. Using equations (8.26) and (8.23) one can show that

$$V_{B|A^M} = \langle X_B^2 \rangle - \langle X_B X_A^M \rangle^2 / \langle (X_A^M)^2 \rangle, \quad (8.56)$$

where using $\langle (X_A^M)^2 \rangle = (V + 1)/2$ and $\langle X_B X_A^M \rangle = \langle X_B X_A \rangle / \sqrt{2}$ finally gives,

$$V_{B|A^M} = T(\chi + 1). \quad (8.57)$$

The secret key rate then reads,

$$K_{\text{RR}} = \frac{1}{2} \log \left[\frac{V_{B|E}}{V_{B|A^M}} \right] \geq \frac{1}{2} \log \left[\frac{1}{V_{B|A} V_{B|A^M}} \right] = \frac{1}{2} \log \left[\frac{1}{T^2(\chi + 1/V)(\chi + 1)} \right], \quad (8.58)$$

as shown in [94].

Implementation of the Optimal Attack

It is trivial to see that in the case of Reverse Reconciliation the entangling cloner is optimal against the protocol based on coherent states and homodyne detection, as Eve information is not changed by Alice measurement. In the case of Direct Reconciliation we see that the only way of minimizing $V_{A^M|E}$ (8.54) is by minimizing $V_{A|E}$ (8.45) which implies that the entangling cloner is also optimal in Direct Reconciliation.

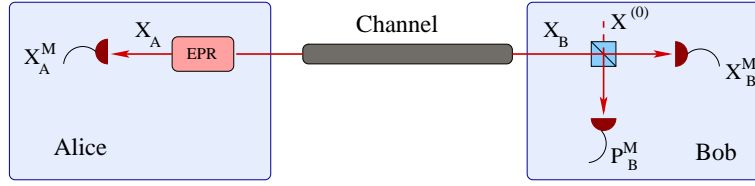


Figure 8.8: Entanglement based scheme of the protocol based on Alice sending squeezed states and Bob applying heterodyne detection. Alice generation of squeezed states is replaced by an EPR state where Alice applies an homodyne detection on one half of the EPR and the other half is sent to Bob.

Squeezed States and Heterodyne Measurement

The protocol based on squeezed states and heterodyne detection is equivalent to an entanglement based scheme where Alice applies an homodyne measurement over mode A and Bob applies heterodyne measurements over mode B , as shown in Fig. 8.8. Restricting to individual attacks and one-way reconciliation, the DR and RR secret key rates read,

$$K_{\text{DR}} = H(A|E) - H(A|B^M), \quad (8.59)$$

$$K_{\text{RR}} = H(B^M|E) - H(B^M|A). \quad (8.60)$$

Similarly as in subsection 8.3, one can show using equation (8.26) that the conditional variance before and after the heterodyne detection are related by,

$$V_{X_{B^M}|Y} = \frac{1}{2} [V_{X_B|Y} + 1], \quad (8.61)$$

for any quadrature Y which is independent of $\hat{X}^{(0)}$ ($\langle \hat{X}^{(0)} Y \rangle = 0$).

Direct Reconciliation

The case of Direct Reconciliation is the symmetric counterpart of the Reverse Reconciliation protocol based in coherent states and homodyne measurement (8.3), as the heterodyne measurement at Bob side affects Alice and Bob mutual but not Eve's information on Alice data. Using equations (8.26) and (8.23) one can show that

$$V_{A|B^M} = \frac{T(V\chi + 1) + V}{T(V + \chi) + 1}. \quad (8.62)$$

The secret key rate then reads,

$$K_{\text{DR}} = \frac{1}{2} \log \left[\frac{V_{A|E}}{V_{A|B^M}} \right] \geq \frac{1}{2} \log \left[\frac{1}{V_{A|B} V_{A|B^M}} \right] = \frac{1}{2} \log \left[\frac{(V + \chi)(T(V + \chi) + 1)}{(V\chi + 1)(T(V\chi + 1) + V)} \right]. \quad (8.63)$$

Reverse Reconciliation

The case of Reverse Reconciliation is the symmetric counterpart of the Direct Reconciliation protocol based in coherent states and homodyne measurement (8.3). By analogy with equations (8.52) we can write,

$$V_{B^M|A} = \frac{1}{2} [V_{B|A} + 1]. \quad (8.64)$$

Eve modes and the vacuum noise at the measurement being not correlated, one can then bound Eve uncertainty using equation (8.35),

$$V_{B^M|E} = \frac{1}{2} [V_{B|E} + 1] \geq \frac{1}{2} \left[\frac{1}{V_{B|A}} + 1 \right]. \quad (8.65)$$

Finally we obtain that the secret key rate reads,

$$K_{RR} = \frac{1}{2} \log \left[\frac{V_{B^M|E}}{V_{B^M|A}} \right] \geq \frac{1}{2} \log \left[\frac{1}{V_{B|A}} \right] = \frac{1}{2} \log \left[\frac{1}{T[\chi + 1/V]} \right]. \quad (8.66)$$

Implementation of the Optimal Attack

The protocol being the symmetric counterpart of the protocol based on Coherent states and homodyne detection, under the exchange of the measurements between Alice and Bob, it is trivial to see that the entangling cloner is again optimal.

The Forgotten Protocol

The protocol based on squeezed states and heterodyne detection has never been studied in the literature before. The most probable reason that explains why none has studied it, is that it is just a noisy version of the protocol based on squeezed and homodyne detection, giving lower secret key rates. Having the disadvantages of a difficult optical implementation of squeezed states and retrieving no gain from applying an heterodyne detection does not seem a very clever idea. Surprisingly when one considers security against collective attacks this protocol becomes interesting, as we will show in the next chapter.

Coherent States and Heterodyne Measurement

The protocol based on coherent states and heterodyne detection [194] is equivalent to an entanglement based scheme where Alice applies an heterodyne measurement over mode A and Bob applies heterodyne measurements over mode B , as shown in Fig. 8.9.

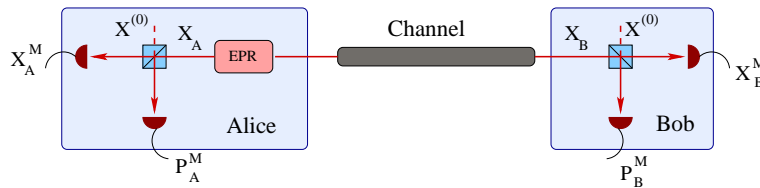


Figure 8.9: Entanglement based scheme of the protocol based on Alice sending coherent states and Bob applying heterodyne detection. Alice generation of coherent is replaced by an EPR state where Alice applies an heterodyne detection on one half of the EPR and the other half is sent to Bob.

Restricting to individual attacks and one-way reconciliation, the DR and RR secret key rates for *each* of the two quadratures read,

$$K_{DR}^{x \text{ or } p} = H(A^M|E) - H(A^M|B^M), \quad (8.67)$$

$$K_{RR}^{x \text{ or } p} = H(B^M|E) - H(B^M|A^M). \quad (8.68)$$

The problem of estimating Bob's uncertainty on Alice's measurements A^M (that is, X_A^M or P_A^M knowing that the other one is also measured) can be reduced to estimating Bob's

uncertainty on each of the quadratures of mode A (X_A, P_A) since Alice's measurements result from mixing mode A with vacuum on a balanced beam splitter, see Fig. 8.9.

Direct Reconciliation

Using equations (8.26) and (8.23) one can show that

$$V_{A^M|B^M} = \frac{1}{2} [V_{A|B^M} + 1] = \frac{1}{2} \left[\frac{(V+1)(T(\chi+1)+1)}{T(V+\chi)+1} \right]. \quad (8.69)$$

Eve uncertainty reads,

$$V_{A^M|E} \geq \frac{1}{2} \left[\frac{1}{V_{A|B}} + 1 \right] = \frac{1}{2} \left[\frac{(V+1)(\chi+1)}{V\chi+1} \right]. \quad (8.70)$$

The secret key rate then reads,

$$K_{\text{DR}} \geq \log \left[\frac{V_{A^M|E}}{V_{A^M|B^M}} \right] = \log \left[\frac{(\chi+1)(T(V+\chi)+1)}{(V\chi+1)(T(\chi+1)+1)} \right]. \quad (8.71)$$

Reverse Reconciliation

Using equations (8.26) and (8.23) one can show that

$$V_{B^M|A^M} = \frac{1}{2} [V_{B^M|A} + 1] = \frac{1}{2} [T(\chi+1) + 1]. \quad (8.72)$$

Eve uncertainty reads,

$$V_{B^M|E} \geq \frac{1}{2} \left[\frac{1}{V_{B|A}} + 1 \right] = \frac{1}{2} \left[\frac{T(V\chi+1)+V}{T(V\chi+1)} \right]. \quad (8.73)$$

The secret key rate then reads,

$$K_{\text{RR}} \geq \log \left[\frac{V_{B^M|E}}{V_{B^M|A^M}} \right] = \log \left[\frac{T(V\chi+1)+V}{T(V\chi+1)(T(\chi+1)+1)} \right]. \quad (8.74)$$

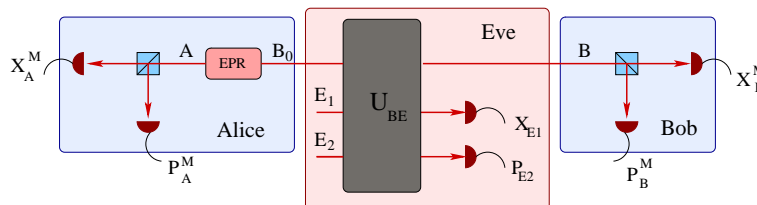
Implementation of the Optimal Attack

The entangling cloner, that is, the optimal attack against the homodyne-based protocols [95], is clearly not optimal here as it allows to extract information about one single quadrature. We may think of adapting it by applying an heterodyne detection on the mode that is entangled with the mode injected in the line (as well as on the output mode of Eve's beamsplitter simulating the losses). However, this is equivalent to having a classical source of noise controlled by Eve, so that the optimal $V_{A(B)|E}$ that Eve can reach coincides with the beamsplitter attack, which does not saturate (8.71) nor (8.74) as the excess noise ϵ only affects Alice and Bob mutual information but does not help Eve to reduce any uncertainty.

Since the time when the heterodyne-based protocol was introduced [194], no attack has been found saturating bounds (8.71) and (8.74). Logically, two possibilities remain open:

1. These bounds are tight but the optimal attacks reaching them remain to be found.
2. These bound are not tight and the (unknown) optimal attacks can not saturate them.

In the next section we will answer this open question by searching the optimal among all possibles attacks.



8.4 No Basis Switching Protocol Optimal Attack

Simplifying the Problem

Thus, the optimal Gaussian attack we seek for corresponds, in the Heisenberg picture, to a symplectic transformation S acting jointly on Alice's mode B_0 and Eve's ancillary modes E_1 and E_2 , that is,

$$[\hat{x}_B, \hat{x}_{E_1}, \hat{x}_{E_2}, \hat{p}_B, \hat{p}_{E_1}, \hat{p}_{E_2}]^T = S [\hat{x}_{B_0}, \hat{x}_{E_1}^{(0)}, \hat{x}_{E_2}^{(0)}, \hat{p}_{B_0}, \hat{p}_{E_1}^{(0)}, \hat{p}_{E_2}^{(0)}]^T, \quad (8.75)$$

where the superscript (0) is used to indicate that the corresponding state is the vacuum. Then, Eve's optimal measurement on her two modes $E \equiv E_1 E_2$ can be assumed to be a homodyne measurement on these two modes in order to estimate either (x_A, p_A) in DR or (x_B, p_B) in RR.

Symmetric Channel without x - p Correlations The symplectic transformation S can be written without loss of generality in a block-diagonal form as

$$S = \begin{pmatrix} S_x & 0 \\ 0 & S_p \end{pmatrix}, \quad (8.76)$$

where S_x and S_p are related by the relation

$$S_p = (S_x^T)^{-1}, \quad (8.77)$$

in order to preserve the canonical commutation relations. Indeed, we start with an initial Gaussian state of covariance matrix $\gamma_{AB_0} \oplus \mathbb{I}_{E_1 E_2}$, which is of the same form as Eq. (8.23). More precisely, it is symmetric in x and p and admits no correlations between x and p . After Eve's Gaussian operation, we have a Gaussian state for modes A and B , which, by Schmidt decomposition, can be purified into a Gaussian 4-mode state by extending the system with modes E_1 and E_2 [105]. This can be understood by applying a symplectic decomposition on modes A and B that converts their joint state into a product of two thermal states. These thermal states can then be written as the reduction of EPR states, shared with Eve's modes E_1 and E_2 . Since this symplectic decomposition does not mix the x and p quadratures, the covariance matrix of the 4-mode pure state is again of the same form as Eq. (8.23). Hence, the symplectic transformation S applied by the eavesdropper does not mix the x and p quadratures. We would like to stress that this form, Eq. (8.76), is not an assumption but rather a simplification originating from the fact that the channels of interest have symmetric uncorrelated noise in x and p , as mentioned above.

The entry of the matrix γ_{AB}^x corresponding to $\langle \hat{x}_B^2 \rangle = T(V + \chi)$ provides constraints on the first row of S_x , since we need to have

$$\hat{x}_B = \sqrt{T}(\hat{x}_{B_0} + \sqrt{\chi} \cos \theta \hat{x}_{E_1}^{(0)} + \sqrt{\chi} \sin \theta \hat{x}_{E_2}^{(0)}), \quad (8.78)$$

where $\theta \in [0, 2\pi]$ is a free parameter. Remember that $\langle \hat{x}_{B_0}^2 \rangle = \langle \hat{x}_A^2 \rangle = V$. Thus, we can write S_x in general as

$$S_x = \sqrt{T} \begin{pmatrix} 1 & \sqrt{\chi} \cos \theta & \sqrt{\chi} \sin \theta \\ a & b & c \\ r & s & t \end{pmatrix}, \quad (8.79)$$

where $\{a, b, c, r, s, t\} \in \mathbb{R}$ are six other free parameters. Using Equation (8.77), we can rewrite S_p as

$$S_p = \frac{1}{d\sqrt{T}} \times \begin{pmatrix} bt - cs & cr - at & as - br \\ \underbrace{\sqrt{\chi}(s \sin \theta - t \cos \theta)}_{r'} & \underbrace{t - r\sqrt{\chi} \sin \theta}_{s'} & \underbrace{r\sqrt{\chi} \cos \theta - s}_{t'} \\ \underbrace{\sqrt{\chi}(c \cos \theta - b \sin \theta)}_{r'} & \underbrace{a\sqrt{\chi} \sin \theta - c}_{s'} & \underbrace{b - a\sqrt{\chi} \cos \theta}_{t'} \end{pmatrix}, \quad (8.80)$$

where $d = \det(S_x)$. Given the symmetry of the channel, the entry of γ_{AB}^p corresponding to $\langle \hat{p}_B^2 \rangle = T(V + \chi)$ provides a constraint on the first row of S_p , in a similar way as for S_x . This yields the three conditions

$$\begin{aligned} bt - cs &= dT, \\ cr - at &= dT\sqrt{\chi} \cos \phi, \\ as - br &= dT\sqrt{\chi} \sin \phi, \end{aligned} \quad (8.81)$$

where $\phi \in [0, 2\pi]$ is a free parameter. Finally, due to the symmetry of the channel in x and p , we consider that Eve's optimal attack gives her the same uncertainty in x and p .

Direct Reconciliation

As before, Eve's uncertainty on Alice's measurements $A^M \equiv (X_A^M, P_A^M)$ can be calculated from the uncertainty of Eve on each of the two quadratures of mode A (X_A, P_A). We have, for example, $V_{X_A^M|X_{E_1}} = \frac{1}{2}(V_{X_A|X_{E_1}} + 1)$, and similarly for the p quadrature. The symmetry of Eve's information on X_A and P_A imposes that

$$V_{X_A|X_{E_1}} = V_{P_A|P_{E_2}} \equiv V_{A|E}. \quad (8.82)$$

Writing the second-order moments of A and E_1 ,

$$\langle \hat{x}_A^2 \rangle = V, \quad (8.83)$$

$$\langle \hat{x}_{E_1}^2 \rangle = T(a^2V + b^2 + c^2), \quad (8.84)$$

$$\langle \hat{x}_A \hat{x}_{E_1} \rangle = a\sqrt{T}\langle \hat{x}_A \hat{x}_{B_0} \rangle = a\sqrt{T(V^2 - 1)}, \quad (8.85)$$

and plugging them into Eq. (8.26), we obtain

$$V_{X_A|X_{E_1}} = \frac{V + \frac{a^2}{b^2+c^2}}{V\frac{a^2}{b^2+c^2} + 1}. \quad (8.86)$$

Similarly, one has for the p quadrature

$$V_{P_A|P_{E_2}} = \frac{V + \frac{r'^2}{s'^2+t'^2}}{V\frac{r'^2}{s'^2+t'^2} + 1}. \quad (8.87)$$

Finally, as a consequence of Eq. (8.82) we can write

$$V_{A|E} = \frac{V + \rho}{V\rho + 1}, \quad (8.88)$$

where

$$\rho \equiv \frac{a^2}{b^2+c^2} = \frac{r'^2}{s'^2+t'^2}. \quad (8.89)$$

Given Eq. (8.78), we see that ρ is proportional to the signal-to-noise ratio of the Alice-to-Eve channel (more precisely, the latter signal-to-noise ratio equals ρV). Thus, by definition, $\rho \geq 0$. Moreover, we can write in analogy with Eq. (8.25) the Heisenberg uncertainty relation

$$V_{X_A|X_{E_1}} V_{P_A|P_{E_2}} \geq 1, \quad (8.90)$$

which, together with Eq. (8.82), implies that $V_{A|E} \geq 1$, or, equivalently, $\rho \leq 1$. Note that the Heisenberg-limited attack in DR corresponds simply to choose $\rho = \chi$.

We will now prove that such a choice is not possible, that is, it is not consistent with the constraints we have on the matrices S_x and S_p . In order to further simplify S_x , we introduce the following change of variables:

$$\begin{aligned} a &= u\sqrt{\rho}, \\ b &= u\sin\xi, \\ c &= u\cos\xi. \end{aligned} \quad (8.91)$$

Using the variables r', s', t' as defined in Eq. (8.80) and the expression of ρ in terms of these variables, Eq. (8.89), we then obtain as explained in Appendix J

$$\left(\frac{\chi - \rho}{\rho}\right) \cos^2(\xi + \theta) = \left(\sin(\xi + \theta) - \sqrt{\rho\chi}\right)^2. \quad (8.92)$$

Using the symmetry of the channel, Eq. (8.81), and the explicit expression of $d = \det S_x$, we obtain a second similar equation (see Appendix J)

$$\left(\frac{\chi - \rho}{\rho}\right) \cos^2(\xi + \theta) = \left(\sin(\xi + \theta) + \frac{1 - T}{T\sqrt{\rho\chi}}\right)^2, \quad (8.93)$$

Expressing the equality between Eqs. (8.92) and (8.93) yields two solutions. The first one, namely $\rho\chi = -(1 - T)/T$, is unphysical since $T \leq 1$, $\rho \geq 0$, and $\chi \geq 0$. The second one yields

$$\sin(\xi + \theta) = \frac{1}{2} \frac{T\chi\rho - (1 - T)}{T\sqrt{\chi\rho}}. \quad (8.94)$$

Furthermore, injecting Eq. (8.94) into Eq. (8.93) gives

$$\cos^2(\xi + \theta) = \left(\frac{1}{2} \frac{T\chi\rho + (1 - T)}{T\sqrt{\chi(\chi - \rho)}}\right)^2. \quad (8.95)$$

Finally, the relation $\cos^2(\xi + \theta) + \sin^2(\xi + \theta) = 1$ provides us with a second-order equation in ρ ,

$$T(T\chi^2 + 4)\rho^2 - 2\chi T(T + 1)\rho + (1 - T)^2 = 0, \quad (8.96)$$

which always admits two solutions for a given channel (i.e. given parameters T and χ),

$$\rho_{\pm} = \frac{\chi T(T + 1) \pm 2\sqrt{T[(T\chi)^2 - (1 - T)^2]}}{T(T\chi^2 + 4)}. \quad (8.97)$$

Looking at Eq. (8.88), we see that minimizing $V_{A|E}$ is equivalent to maximizing ρ , that is, choosing ρ_+ . Thus, Eve's minimum uncertainty on Alice's measurement reads,

$$V_{A^M|E}^{\min} = \frac{1}{2} [V_{A|E}^{\min} + 1] = \frac{1}{2} \frac{(V + 1)(\rho_+ + 1)}{V\rho_+ + 1}, \quad (8.98)$$

and the lower bound on the DR secret key rate reads

$$\begin{aligned} K_{\text{DR}} &= \log \left[\frac{V_{A^M|E}^{\min}}{V_{A^M|B^M}} \right] \\ &= \log \left[\frac{(\rho_+ + 1)(T(V + \chi) + 1)}{(V\rho_+ + 1)(T(\chi + 1) + 1)} \right]. \end{aligned} \quad (8.99)$$

Interestingly, Eq. (8.98) is similar to its counterpart for the Heisenberg-limited attack, Eq. (8.70), but with ρ_+ replacing χ . It can easily be checked that $\rho_+ < \chi$, so that the highest possible signal-to-noise ratio of the Alice-to-Eve channel is strictly lower than the one deduced from Heisenberg uncertainty relations. Hence, Eve's optimal attack is less powerful than expected from Heisenberg relations.

This is illustrated in Fig. 8.11, where the secret key rates have been plotted for experimental realistic values of V and ϵ . The lower bound deduced from the Heisenberg relations is satisfied, but loose with respect to the actual key rate.

Reverse Reconciliation

Combining Eqs. (8.75) and (8.79), we obtain the second-order moments of B and E_1

$$\langle \hat{x}_B^2 \rangle = T(V + \chi), \quad (8.100)$$

$$\langle \hat{x}_{E_1}^2 \rangle = T(a^2V + b^2 + c^2), \quad (8.101)$$

$$\langle \hat{x}_B \hat{x}_{E_1} \rangle = T(aV + b\sqrt{\chi} \cos \theta + c\sqrt{\chi} \sin \theta). \quad (8.102)$$

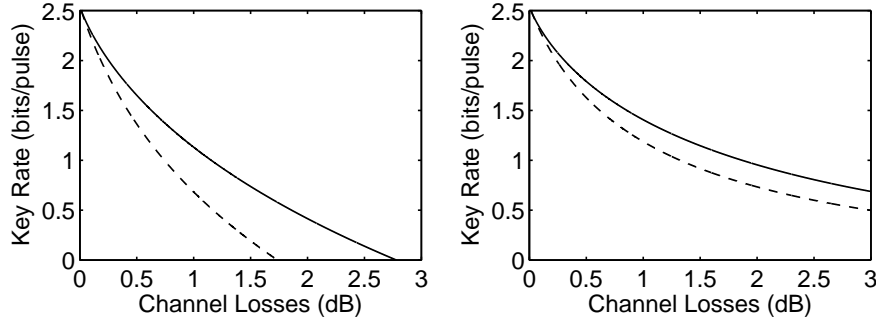


Figure 8.11: Secret key rate as a function of the line losses for the optimal (solid line) and Heisenberg-limited (dashed line) attack. The curves are plotted for experimentally realistic values, $V = 12$ and $\epsilon = 0.01$, in direct reconciliation (left panel) or reverse reconciliation (right panel).

This results in

$$V_{X_B|X_{E1}} = T \frac{\left[\frac{b^2+c^2}{a^2} + \chi - \frac{2\sqrt{\chi}}{a}(b \cos \theta + c \sin \theta) \right] V + \frac{\chi}{a^2}(b \sin \theta - c \cos \theta)^2}{V + \frac{b^2+c^2}{a^2}}, \quad (8.103)$$

where we have used Eq. (8.26). Similarly, using the symmetry of the channel, Eq. (8.81), we can write,

$$V_{P_B|P_{E2}} = T \frac{\left[\frac{s'^2+t'^2}{r'^2} + \chi - \frac{2\sqrt{\chi}}{r'}(s' \cos \phi + t' \sin \phi) \right] V + \frac{\chi}{r'^2}(s' \sin \phi - t' \cos \phi)^2}{V + \frac{s'^2+t'^2}{r'^2}}. \quad (8.104)$$

Imposing the symmetry of Eve's information on X_B and P_B in analogy with Eq. (8.82), that is,

$$V_{X_B|X_{E1}} = V_{P_B|P_{E2}} \equiv V_{B|E}, \quad (8.105)$$

gives the three conditions

$$\frac{r'^2}{s'^2+t'^2} = \frac{a^2}{b^2+c^2} = \rho, \quad (8.106)$$

$$\frac{s' \cos \phi + t' \sin \phi}{r'} = \frac{b \cos \theta + c \sin \theta}{a} = \frac{\sin(\xi + \theta)}{\sqrt{\rho}}, \quad (8.107)$$

$$\frac{s' \sin \phi - t' \cos \phi}{r'} = \frac{b \sin \theta - c \cos \theta}{a} = \frac{\cos(\xi + \theta)}{\sqrt{\rho}}. \quad (8.108)$$

Note that condition (8.106) is exactly the same as in direct reconciliation. Surprisingly, it so happens that this condition is sufficient to find an expression for $V_{B|E}$ which is the same as in direct reconciliation, making it unnecessary to use the other two conditions. Indeed, Eve's uncertainty on the quadratures of mode B can be rewritten as

$$V_{B|E} = T \frac{[1 + \chi\rho - 2\sqrt{\chi\rho}\sin(\xi + \theta)]V + \chi\cos^2(\xi + \theta)}{V\rho + 1}. \quad (8.109)$$

Then, using the definition of $\sin(\xi + \theta)$ coming from Eq. (8.94) as well as Eq. (8.96), we obtain

$$\cos^2(\xi + \theta) = \frac{\rho}{T\chi}, \quad (8.110)$$

$$1 + \chi\rho - 2\sqrt{\chi\rho}\sin(\xi + \theta) = 1/T, \quad (8.111)$$

which gives $V_{B|E} = V_{A|E}$. Therefore, just like in direct reconciliation, Eve's uncertainty on the quadratures of mode B is minimized by choosing ρ_+ ,

$$V_{B|E}^{\min} = \frac{V + \rho_+}{V\rho_+ + 1}. \quad (8.112)$$

Then, Eve's uncertainty on Bob's measured values becomes

$$V_{B^M|E}^{\min} = \frac{1}{2} [V_{B|E}^{\min} + 1] = \frac{1}{2} \frac{(V + 1)(\rho_+ + 1)}{V\rho_+ + 1}, \quad (8.113)$$

so that the RR secret key rate reads

$$\begin{aligned} K_{\text{RR}} &= \log \left[\frac{V_{B^M|E}^{\min}}{V_{B^M|A^M}} \right] \\ &= \log \left[\frac{(V + 1)(\rho_+ + 1)}{(V\rho_+ + 1)(T(\chi + 1) + 1)} \right]. \end{aligned} \quad (8.114)$$

This rate is illustrated in Fig. 8.11, where it is compared with the lower bound deduced from the Heisenberg relations in RR. We conclude again that the Heisenberg-limited attack is not reachable.

8.5 Optical Setup Achieving the Optimal Attack

In Section 8.4, we have reduced the problem of maximizing Eve's information to that of optimizing a single parameter ρ , the other parameters remaining free. This implies that the optical implementation of the best Gaussian attack is not unique. In this Section, we present two particularly interesting examples of such an optical implementation, namely the teleportation attack and the “feed-forward” attack. Note that the latter attack was also considered in Ref. [194], where it was noticed that it curiously does not reach the Heisenberg limit.

Teleportation Attack

The teleportation attack consists in Eve applying a continuous-variable quantum teleportation where the input is Alice's outgoing mode and the output is given to Bob, as shown in Fig. 8.12. Eve extracts information from the outcomes (X_E^M, P_E^M) of her Bell measurement performed on Alice's outgoing mode B_0 together with one of the modes (E_1') of an EPR state. It is easy to see that there are two limiting cases. If the squeezing factor r of the EPR pair is zero, implying that E_1' is in a vacuum state, then the scheme becomes equivalent to an heterodyne measurement of B_0 by Eve followed by the classical preparation of a coherent state (the vacuum state in mode E_2' which is displaced by some amount depending on X_E^M and P_E^M). This situation corresponds to an entanglement-breaking channel giving no secret key. On the contrary, if the squeezing factor r is infinite, the teleportation succeeds perfectly and Eve gets no information at all due to the infinite noise in the thermal state E_1' . This situation corresponds to a perfect channel with no losses and no excess noise ($T = 1, \epsilon = 0$). We will now show that for any intermediate value of r , such a teleportation attack can be made optimal.

Since all the involved canonical transformations are symmetric in x and p , we will detail the proof for the x quadrature only. Eve starts by preparing two squeezed vacuum states, mode E_2 squeezed along x and mode E_1 squeezed along p ,

$$\hat{x}_1 \equiv \hat{x}_{E_1} = e^r \hat{x}_1^{(0)}, \quad (8.115)$$

$$\hat{x}_2 \equiv \hat{x}_{E_2} = e^{-r} \hat{x}_2^{(0)}, \quad (8.116)$$

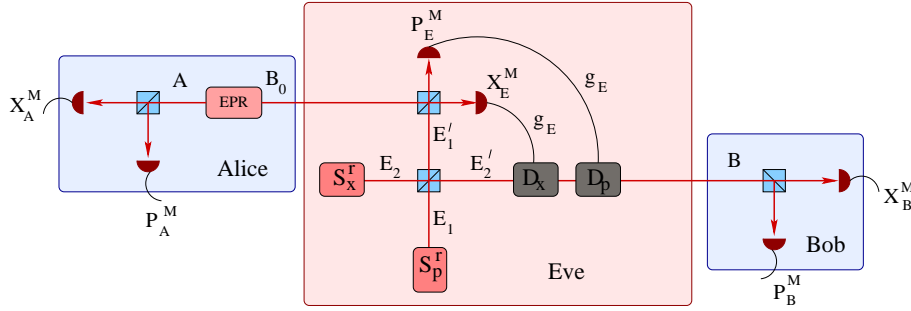


Figure 8.12: Teleportation attack against the (entanglement-based scheme of the) Gaussian protocol based on Alice sending coherent states and Bob applying heterodyne detection. Eve first generates an EPR pair (E_1', E_2') by mixing a x -squeezed vacuum state (E_2) with a p -squeezed vacuum state (E_1) at a balanced beamsplitter. Then, she performs a Bell measurement on Alice's outgoing mode B_0 together with E_1' . Depending on the measurement outcome and the fixed gain g_E , she then displaces mode E_2' by x (D_x) and p (D_p). The resulting state is sent to Bob. By tuning the squeezing parameter r and the gain g_E , Eve can simulate any Gaussian channel (T, χ) and extract the optimal amount of information.

where we omitted the subscript E to lighten the notations. Subsequently we mix them on a balanced beamsplitter, thereby generating an EPR state

$$\hat{x}'_1 = [e^{-r}\hat{x}_2^{(0)} - e^r\hat{x}_1^{(0)}]/\sqrt{2}, \quad (8.117)$$

$$\hat{x}'_2 = [e^{-r}\hat{x}_2^{(0)} + e^r\hat{x}_1^{(0)}]/\sqrt{2}. \quad (8.118)$$

Eve then applies a Bell measurement by mixing E_1' and B_0 on a balanced beamsplitter, and measuring x on one output and p on the other,

$$\hat{x}_E^M = \frac{1}{\sqrt{2}}[\hat{x}_{B_0} + \hat{x}'_1] = \frac{1}{\sqrt{2}}\hat{x}_{B_0} + \frac{1}{2}[e^{-r}\hat{x}_2^{(0)} - e^r\hat{x}_1^{(0)}]. \quad (8.119)$$

Next, Eve displaces her mode E_2' by an amount proportional to the measurement outcome X_E^M (multiplied by the classical gain g_E) and sends it to Bob, giving

$$\begin{aligned} \hat{x}_B &= \hat{x}'_2 + g_E \hat{x}_E^M \\ &= \frac{g_E}{\sqrt{2}}\hat{x}_{B_0} + \frac{e^r}{\sqrt{2}}\left[1 - \frac{g_E}{\sqrt{2}}\right]\hat{x}_1^{(0)} + \frac{e^{-r}}{\sqrt{2}}\left[1 + \frac{g_E}{\sqrt{2}}\right]\hat{x}_2^{(0)}. \end{aligned} \quad (8.120)$$

In order to comply with $\langle \hat{x}_B^2 \rangle = T(V + \chi)$, we need to fix g_E and r in such a way that

$$g_E = \sqrt{2T}, \quad (8.121)$$

$$T\chi = (1 + T) \cosh 2r - 2\sqrt{T} \sinh 2r. \quad (8.122)$$

Direct Reconciliation.

Writing the second-order moments of \hat{x}_A and \hat{x}_E^M , namely

$$\langle \hat{x}_A^2 \rangle = V, \quad (8.123)$$

$$\langle (\hat{x}_E^M)^2 \rangle = (V + \cosh 2r)/2, \quad (8.124)$$

$$\langle \hat{x}_A \hat{x}_E \rangle = \langle \hat{x}_A \hat{x}_{B_0} \rangle / \sqrt{2} = \sqrt{(V^2 - 1)/2}, \quad (8.125)$$

one can show, using Eq. (8.26), that Eve's uncertainty on Alice's data is

$$V_{A|E} = \frac{V \cosh 2r + 1}{V + \cosh 2r}. \quad (8.126)$$

By choosing

$$\rho = \frac{1}{\cosh 2r}, \quad (8.127)$$

this expression for $V_{A|E}$ coincides with Eq. (8.88). Combining Eq. (8.122) with the relation $\cosh^2 2r - \sinh^2 2r = 1$, we see that ρ must satisfy the second-order polynomial equation (8.96), whose solution gives the value of ρ that optimizes Eve's information. Equation (8.96) having two possible solutions ρ_{\pm} generating the same quantum channel (T, χ) , we then have two possible solutions for the squeezing parameter r . Looking at Eq. (8.127), we see that the squeezing parameter corresponding to the optimal choice ρ_+ is the lowest of the two solutions since it corresponds to the minimum added noise on Eve's measurement.

Reverse Reconciliation.

Using Eqs. (8.26), (8.122), (8.124), and

$$\langle \hat{x}_B \hat{x}_E^M \rangle = \frac{1}{\sqrt{2}} [V\sqrt{T} - \sinh 2r + \sqrt{T} \cosh 2r], \quad (8.128)$$

one can show that Eve's uncertainty on each of Bob's quadratures reads

$$V_{B|E} = \frac{V \cosh 2r + 1}{V + \cosh 2r} = V_{A|E}, \quad (8.129)$$

implying that the teleportation attack is also optimal (choosing the lowest squeezing parameter) for the reverse reconciliation protocol.

Feed-forward Attack

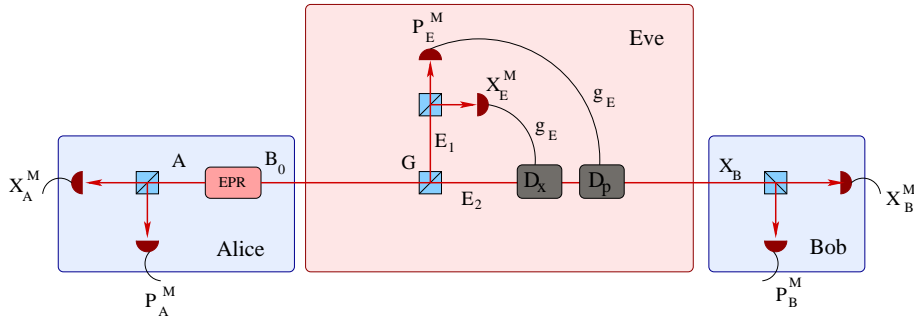


Figure 8.13: Entanglement based scheme of Eve “feed-forward” attack over the protocol based on Alice sending coherent states and Bob applying heterodyne detection. Eve extracts part of the signal sent by Alice using a beamsplitter (transmittance G) and applies an heterodyne detection on it (over mode E_1). Depending on the measurement result times a given fixed gain g_E Eve displaces mode E_2 over x (D_x) and p (D_p). The resulting state is then sent to Bob. By tuning the transmittance of the beamsplitter (G) and the gain (g_E) Eve can simulate any Gaussian channel (T, χ) and extract the optimal amount of information.

In the case of a noisy channel with no losses ($T = 1$) and direct reconciliation, Eve's optimal teleportation attack is exactly the same scheme as the one proposed in Ref. [5] to reach an optimal tradeoff between disturbance and state estimation for coherent states (when the success of both processes is measured using the fidelity). This is not surprising since optimally estimating the coherent state sent by Alice while minimizing its disturbance is exactly what Eve attempts to achieve in her optimal attack in direct reconciliation. In Ref. [5], two alternative schemes to the teleportation reaching the same optimal tradeoff were also presented, the “feed-forward” attack and the asymmetric cloning machine. Those two schemes can very naturally be extended to our case ($T \leq 1$) if we allow for different mean values for the input and output modes, which gives rise to new optical schemes for the optimal attack.

For example, it can be checked that Eve can realize an optimal attack (both in DR and RR) using the “feed-forward” scheme described in Fig. 8.13 by fixing the parameters of the beamsplitter transmittance G and the feed-forward gain g_E as

$$G = \frac{1 - \rho_+}{1 + \rho_+}, \quad (8.130)$$

$$g_E = (\sqrt{T} - \sqrt{G}) \sqrt{\frac{2}{1 - G}}. \quad (8.131)$$

8.6 Security Analysis of the Gaussian Protocols

For a lossy channel (no excess noise, $\epsilon = 0$), we observe in Fig. 8.14 that the protocol based on squeezed states and homodyne detection gives the highest rate together with the no basis switching protocol in both Direct and Reverse Reconciliation. The

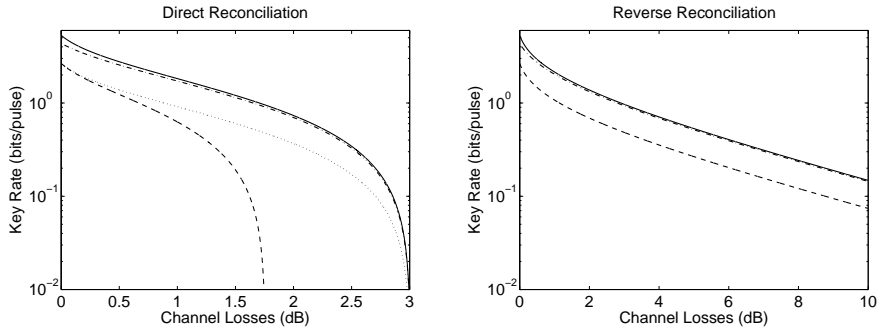


Figure 8.14: Secret key rate as a function of the channel losses (measured in dB) for a lossy channel ($\epsilon = 0$) for all the Gaussian protocols: squeezed states and homodyne detection (solid line), coherent states and homodyne detection (dotted line), coherent states and heterodyne detection (dot-dashed line) and squeezed states and heterodyne detection (dashed line). The curves are plotted for experimental realistic modulation $V = 40$.

worst protocol in Direct Reconciliation is the *forgotten protocol* based on Alice sending squeezed states and Bob applying heterodyne detection. In the case of Reverse Reconciliation there are two protocols giving the worst secret rate, the protocol based on coherent states and homodyne measurement and the one based on squeezed states and heterodyning, as we can easily check using equations (8.58) and (8.66) and realizing that $T(\chi + 1) = 1$ for $\epsilon = 0$.

Once we move to a noisy channel we observe that the protocol based on squeezed states and homodyne detection remains optimal but the no basis switching protocol

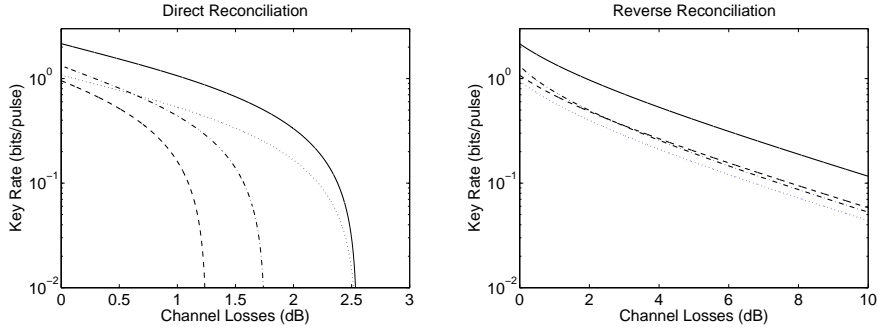


Figure 8.15: Secret key rate as a function of the channel losses (measured in dB) for a noisy channel ($\epsilon = 0.2$), for all the Gaussian protocols: squeezed states and homodyne detection (solid line), coherent states and homodyne detection (dotted line), coherent states and heterodyne detection (dot-dashed line) and squeezed states and heterodyne detection (dashed line). The curves are plotted for experimental realistic modulation $V = 40$.

becomes very sensitive to the noise, as shown in Fig. 8.15. After a given threshold, it starts to perform worse than the protocol based on coherent states and homodyning in DR and than squeezed states and heterodyne detection in RR.

Tolerable Excess Noise

In Direct Reconciliation, for both implementations based on Alice using coherent states, the heterodyne-based protocol gives an advantage over the homodyne-based protocol only for line losses below some threshold. This threshold can be shown to decrease for increasing ϵ , so that the maximum tolerable excess noise is actually higher for the homodyne-based protocol, as we can see in Fig. 8.16 where we plot the maximal excess noise tolerable by each protocol as a function of the channel loss.

In Reverse Reconciliation we observe a symmetric effect for the protocol using heterodyne detection, where the no-switching protocol gives an advantage over the protocol based on squeezed states and heterodyne detection only for line losses below some threshold, giving a lower tolerable excess noise as shown in Fig. 8.17.

In Fig. 8.18 we plot the tolerance to excess noise of all the Gaussian one-way protocols against individual attacks by Eve. Where we observe a relevant behavior of continuous variables QKD, the difference existing between Direct and Reverse Reconciliation maximal ranges. We remark that the range in Direct Reconciliation is limited to 3dB while there is no theoretical limitation in the case of Reverse Reconciliation.

The coherent states and homodyne detection secret key rate in Direct Reconciliation (8.55) being up to a constant factor exactly the same as that of the protocol based on squeezed states and homodyne detection (8.34), it is then trivial to see that they give the same tolerable excess noise as they become null for the same excess noise. Similarly, the squeezed state and heterodyning protocol gives the same tolerable excess noise (8.66) as that of the protocol based on squeezed states and homodyne detection (8.36) in Reverse Reconciliation.

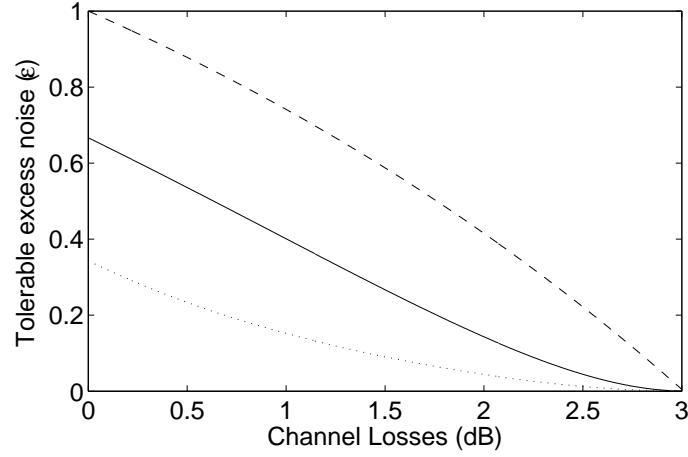


Figure 8.16: Tolerable excess noise ϵ as a function of the channel losses (measured in dB) for high modulation ($V = 1000$) and DR for the two protocols based on coherent states; coherent states with heterodyne detection (solid line) and homodyne detection (dashed line). The dotted line marks the transition where the key rate of the two protocols is equivalent.

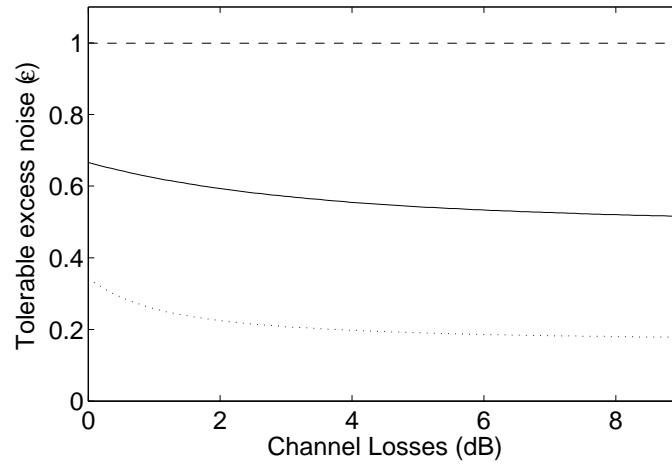


Figure 8.17: Tolerable excess noise ϵ as a function of the channel losses (measured in dB) for high modulation ($V = 1000$) and RR for the two protocols based on heterodyne detection; coherent states with heterodyne detection (solid line) and squeezed states and heterodyne detection (dashed line). The dotted line shows the channels (couple T, ϵ) that give the same secret key rate for both protocols.

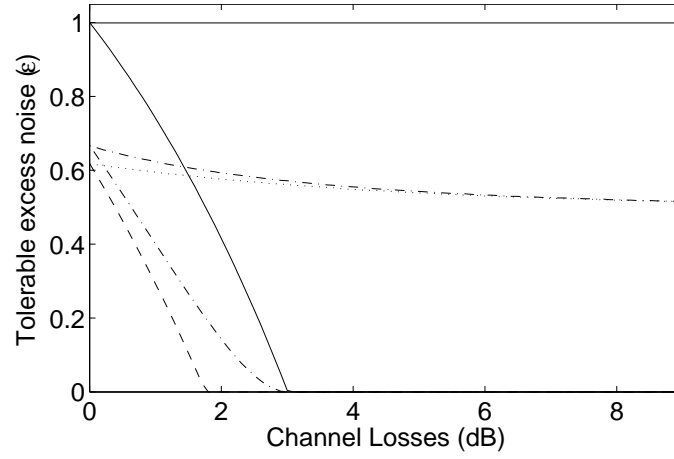


Figure 8.18: Tolerable excess noise ϵ as a function of the channel losses (measured in dB) for high modulation ($V = 1000$) for all the Gaussian protocols: squeezed states and homodyne detection (solid line), coherent states and homodyne detection (dotted line), coherent states and heterodyne detection (dot-dashed line) and squeezed states and heterodyne detection (dashed line). The curves vanishing at (or above) 3dB correspond to DR, whereas the rest refer to RR.

Chapter 9

CV-QKD: Collective Attacks

9.1 Introduction

In this chapter we are going to study the security of the continuous variable quantum key distribution (CV-QKD) protocols based on Gaussian modulation of Gaussian states presented in Chapter 7. All the protocols can be described in an unified way using an entanglement-based description of CV-QKD, as shown in Fig. 9.1, where Alice and Bob distill a secret key from their measurement results a and b respectively. In

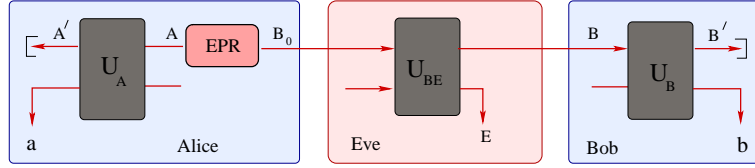


Figure 9.1: Entanglement-based scheme for CV-QKD. Alice's preparation is modelled by a measurement U_A on her half of an EPR pair. The channel is modelled by an unitary interaction between mode B and Eve ancilla's E . Finally, Bob's measurement is modelled by U_B .

realistic protocols, Alice and Bob do not achieve the Holevo bound, but only extract the mutual information $I_{ab} = S(a:b)$. In contrast, Eve is assumed to have no technological limitation, so, by collective attacks, she can attain the Holevo bound $\chi_{aE} = S(a:E)$. Then, using our notation, the achievable secret key rate reads [155],

$$K_{DR} = S(a:b) - S(a:E), \quad (9.1)$$

for Direct Reconciliation (DR) and

$$K_{RR} = S(a:b) - S(b:E), \quad (9.2)$$

for Reverse Reconciliation (RR), where the function $K_{DR(RR)}(\rho_{AB})$ depends on the choice of the measurements done by Alice and Bob (and on the sifting if any), but does not depend on the purification of ρ_{AB} .

9.2 Optimality of Gaussian Collective Attacks

The proof of optimality of Gaussian attacks among the family of collective attacks is strikingly similar to that of individual attacks presented in Chapter 8, using the

entanglement-based version of CV-QKD supplemented by the physical model of measurement and the optimality of Gaussian states derived in [198].

Before presenting the proof we briefly remind the results of [198], where it is shown that for a given function f which is: (i) continuous in trace norm; (ii) invariant under local "Gaussification" operation U_G ; (iii) strongly sub-additive. Then, for every bipartite state ρ_{AB} with covariance matrix γ_{AB} , we have that $f(\rho_{AB}) \leq f(\rho_{AB}^G)$ where ρ_{AB}^G is the Gaussian state with the same γ_{AB} .

In the entanglement-based Gaussian protocols presented in Chapter 7 Alice and Bob were using either homodyne or heterodyne detection. The next optimality result is true for more general POVM measurements, such as noisy homodyne measurements. The only constraint on the measurements is that they must commute with the "Gaussification" operation, which is satisfied if the measurement acts independently on x and p quadratures.

Proof

Alice and Bob mutual information $S(a:b)$ being fixed by the data obtained by the two partners and the efficiency of the reconciliation, in order to prove the optimality of Gaussian attacks it is then enough to prove that Eve information $S(a:E)$ is maximized when she applies a Gaussian map. For this, we need to use the extension of this function over $2N$ modes ($\mathbf{A} = A_1 A_2 \dots A_N$, $\mathbf{B} = B_1 B_2 \dots B_N$), $S(\mathbf{a}:E)$. Remark that E being the purification of \mathbf{AB} the quantum mutual information $S(\mathbf{A}:E)$ is a function of $\rho_{\mathbf{AB}}$. Note that $S(\mathbf{a}:E)$ restricts to $S(a:E)$ when $N = 1$.

Continuity If $\|\rho_{\mathbf{AB}}^{(n)} - \rho_{\mathbf{AB}}\|_1 \leq \epsilon$, using Uhlmann's theorem and the well-known relations between the fidelity and trace distance (see Appendix I), we can find a purification $|\Psi\rangle_{\mathbf{AB}E}^{(n)}$ ($|\Psi\rangle_{\mathbf{AB}E}$) of $\rho_{\mathbf{AB}}^{(n)}$ ($\rho_{\mathbf{AB}}$) such that $\|\hat{\Psi}_{\mathbf{AB}E}^{(n)} - \hat{\Psi}_{\mathbf{AB}E}\|_1 \leq 2\sqrt{\epsilon}$. Then, considering that partial trace can only decrease the trace norm [136], we have $\|\rho_{\mathbf{a}E}^{(n)} - \rho_{\mathbf{a}E}\|_1 \leq 2\sqrt{\epsilon}$, $\|\rho_E^{(n)} - \rho_E\|_1 \leq 2\sqrt{\epsilon}$ and $\|\rho_{\mathbf{a}}^{(n)} - \rho_{\mathbf{a}}\|_1 \leq 2\sqrt{\epsilon}$. Finally, the continuity of von Neumann entropies implies the continuity of $S(\mathbf{a}:E)$. \square

Invariance Under Local Gaussification Unitaries Applying the local Gaussification operation $U_G \otimes U_G$ on the product states $|\psi\rangle_{ABE}^{\otimes N}$ (as shown in Fig. 9.2 for $N = 2$), we obtain the state $|\tilde{\psi}\rangle_{ABE}$. After the measurements on Alice's and Bob's

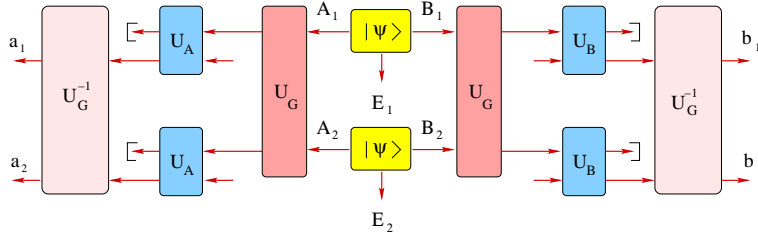


Figure 9.2: Invariance under local "Gaussification" unitaries: U_G can be interchanged with the measurement U_A , then U_G^{-1} and U_G cancel each other.

sides, the state becomes $\tilde{\rho}_{\mathbf{a}bE}$. But because the measurement does not mix the x and p quadratures and neither does the Gaussification operation, the two can be interchanged, by applying $U_G^\dagger \otimes U_G^\dagger$ on modes \mathbf{a} and \mathbf{b} we recover the state $\rho_{\mathbf{a}bE}^{\otimes N}$, which coincides with the state obtained by directly measuring $|\psi\rangle_{ABE}^{\otimes N}$ without Gaussification. Since the two states $\tilde{\rho}_{\mathbf{a}bE}$ and $\rho_{\mathbf{a}bE}^{\otimes N}$ are related by a local unitary operation $U_G^\dagger \otimes \mathbb{I}$ and

since the von Neumann entropies appearing in $S(\mathbf{a}:E)$ are invariant under (any) local unitaries, we obtain the invariance of $S(\mathbf{a}:E)$ under local Gaussification unitaries. \square

Strong Subadditivity We will restrict the proof to two modes on each side, $A_{1,2}$ and $B_{1,2}$, as shown in Fig. 9.3, where E is the purification of $A_{1,2}B_{1,2}$. The generalization to $N > 2$ being straightforward. Using the definition of the mutual entropy,

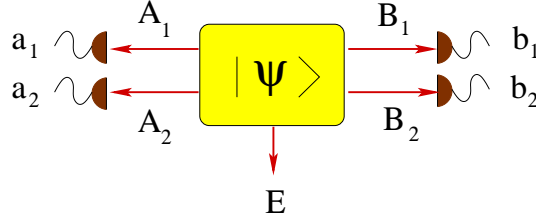


Figure 9.3: Alice and Bob share a quantum state $A_{1,2}B_{1,2}$ where Alice applies independent measurements on each mode $M_{A_1} \otimes M_{A_2}$, and so does Bob $M_{B_1} \otimes M_{B_2}$. Eve holds the purification (E) of $A_{1,2}B_{1,2}$. Remark that the purification of A_1B_1 (A_2B_2) is A_2B_2E (A_1B_1E).

$$\begin{aligned}
 S(a_1, a_2; E) &= S(a_1, a_2) - S(a_1, a_2|E) \\
 &= \underbrace{S(a_1, a_2)}_{\leq S(a_1) + S(a_2)} - [\underbrace{S(a_1|a_2E)}_{\geq S(a_1|A_2B_2E)} + \underbrace{S(a_2|a_1E)}_{\geq S(a_2|A_1B_1E)} + \underbrace{S(a_1:a_2|E)}_{\geq 0}] \\
 &\leq S(a_1) + S(a_2) - S(a_1|A_1B_1E) - S(a_2|A_2B_2E), \tag{9.3}
 \end{aligned}$$

where we used the subadditivity, the strong subadditivity of the entropy and "conditioning does not increase entropy"¹. Finally, noticing that the purification of A_1B_1 (A_2B_2) is $E_1 = A_2B_2E$ ($E_2 = A_1B_1E$) we obtain

$$S(a_1, a_2) \leq S(a_1:E_1) + S(a_2:E_2). \tag{9.4}$$

The additivity being a straightforward consequence of the additivity of von Neumann entropies. \square

Thus, we have proved that for all bipartite quantum states ρ_{AB} with covariance matrix γ_{AB} , one has $K_{DR}(\rho_{AB}) \geq K_{DR}(\rho_{AB}^G)$. This means that $K_{DR}(\rho_{AB}^G)$ is a lower bound on the secret key rate for any protocol (even non-Gaussian) and collective attack (including non-Gaussian). The only requirement for this result to hold is that Alice and Bob use the second-order moments of the quadratures in order to calculate this bound. In particular, for the Gaussian-modulation protocols of [44, 94, 96, 194], Eve's optimal attack is a Gaussian attack, in which case the bound is saturated. Note that the above proof concerns DR, see Eq. (9.1), but its extension to RR Eq. (9.2) is straightforward: one simply needs to interchange $a \leftrightarrow b$ and $A \leftrightarrow B$.

9.3 Security Analysis of Gaussian Protocols

Importantly, this bound $K_{DR(RR)}(\rho_{AB}^G)$ can easily be computed from the observed data since one simply needs to calculate the entropy of thermal states. In the following we are going to analyse the security of all existing Gaussian protocols.

¹In Chapter 5 and 6 we proved that "conditioning does not increase entropy" and strong subadditivity are strictly equivalent.

Squeezed States and Homodyne Detection

The protocol based on squeezed states and homodyne measurement [44] is equivalent to an entanglement based scheme where Alice and Bob apply homodyne measurements over modes A and B respectively, as shown in Fig. 9.4. The quantum state ρ_{AB} before

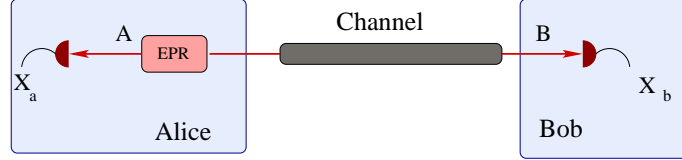


Figure 9.4: Entanglement based scheme of the protocol based on Alice sending squeezed states and Bob applying homodyne detection. Alice generation of squeezed states is replaced by an EPR state where Alice applies an homodyne detection on one half of the EPR and the other half is sent to Bob.

Alice and Bob measurements is a Gaussian two mode state with null mean value and covariance matrix,

$$\gamma_{AB} = \begin{pmatrix} a\mathbb{I} & c\sigma_z \\ c\sigma_z & b\mathbb{I} \end{pmatrix} = \begin{pmatrix} V\mathbb{I} & \sqrt{T(V^2-1)}\sigma_z \\ \sqrt{T(V^2-1)}\sigma_z & T(V+\chi)\mathbb{I} \end{pmatrix}, \quad (9.5)$$

where σ_z reads

$$\sigma_z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad (9.6)$$

V is the variance of Alice output thermal state and T and $\chi = (1-T)/T + \epsilon$ are respectively the transmittance and noise referred to the input of the Gaussian channel (ϵ being the excess noise referred to the input). The secret key rate reads,

$$K = S(x_A:x_B) - I_E, \quad (9.7)$$

where Eve information I_E reads, $S(x_A:E)$ for DR and $S(x_B:E)$ for RR.

Alice and Bob Mutual Information

Alice and Bob mutual information

$$S(x_A:x_B) = S(x_A) - S(x_A|x_B), \quad (9.8)$$

can be calculated using the techniques developed in chapter 8,

$$S(x_A:x_B) = \frac{1}{2} \log \left[\frac{V_A}{V_{A|B}} \right] = \frac{1}{2} \log \left[\frac{V+\chi}{\chi+1/V} \right], \quad (9.9)$$

which is the same for Direct and Reverse Reconciliation.

Eve Information: Direct Reconciliation

Eve quantum information on Alice measurement

$$S(x_A:E) = S(E) - S(E|x_A), \quad (9.10)$$

can be calculated using the following technique. First we use the fact that system E purifies AB so we can write $S(E) = S(AB)$. Secondly, after Alice projective measurement X_A the system BE being pure, we have $S(E|x_A) = S(B|x_A)$. For Gaussian

states $S(B|x_A)$ is the same for all x_A , being just a function of the covariance matrix γ_{AB} . $S(AB)$ is a function of the symplectic eigenvalues $\lambda_{1,2}$ of γ_{AB} which reads

$$S(AB) = G[(\lambda_1 - 1)/2] + G[(\lambda_2 - 1)/2], \quad (9.11)$$

where

$$G(x) = (x + 1) \log(x + 1) - x \log x, \quad (9.12)$$

is the Von Neumann entropy of a thermal state (calculated in Chapter 6) and

$$\lambda_{1,2}^2 = \frac{1}{2} \left[\Delta \pm \sqrt{\Delta^2 - 4D^2} \right], \quad (9.13)$$

where we have used the notations

$$\begin{aligned} \Delta &= a^2 + b^2 - 2c^2 = V^2(1 - 2T) + 2T + T^2(V + \chi)^2, \\ D &= ab - c^2 = T(V\chi + 1). \end{aligned} \quad (9.14)$$

$S(B|x_A) = G[(\lambda_3 - 1)/2]$ is a function of the symplectic eigenvalue λ_3 of the covariance matrix $\gamma_B^{x_a}$ of Bob mode after Alice projective measurement (using equation (2.57)),

$$\gamma_B^{x_a} = \gamma_B - \sigma_{AB}^T (X \gamma_A X)^{MP} \sigma_{AB}, \quad (9.15)$$

which gives,

$$\gamma_B^{x_a} = \begin{pmatrix} b - c^2/a & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} T(\chi + 1/V) & 0 \\ 0 & T(V + \chi) \end{pmatrix}. \quad (9.16)$$

The square of symplectic eigenvalue λ_3 reads,

$$\lambda_3^2 = b(b - c^2/a) = T^2(V + \chi)(\chi + 1/V). \quad (9.17)$$

Eve Information: Reverse Reconciliation

Eve quantum information on Bob measurement can be calculated in a similar way as for Direct Reconciliation,

$$S(x_B:E) = S(E) - S(E|x_B). \quad (9.18)$$

Because of the symmetry of the entanglement-based scheme we just need to interchange $a \leftrightarrow b$ in equation (9.16), in order to calculate $\gamma_A^{x_b}$. The square of the symplectic eigenvalue λ_3 of the covariance matrix $\gamma_A^{x_b}$ then reads,

$$\lambda_3^2 = a(a - c^2/b) = V \frac{V\chi + 1}{V + \chi}. \quad (9.19)$$

Coherent States and Homodyne Detection

The protocol based on coherent states and homodyne measurement [96] is equivalent to an entanglement based scheme where Alice applies an heterodyne measurement on mode A , as shown in Fig. 9.5, where the heterodyne measurement is modeled by combining mode A and a vacuum ancilla C in a balanced beamsplitter and measuring x on A and p on C . Bob continues to apply homodyne measurements over mode B , obliging Alice to drop one of both measurement results during the sifting step. The secret key rate then reads,

$$K = S(x_A^M : x_B) - I_E, \quad (9.20)$$

where Eve information I_E is $S(x_A^M : E)$ for DR and $S(x_B : E)$ for RR, Note that we use the variable x_A^M here (not A), since in this protocol Alice does not measure one single quadrature but a pair of conjugate quadratures [x_A^M stands for the measurement of x quadrature of mode A , given that the conjugate quadrature p is simultaneously measured].

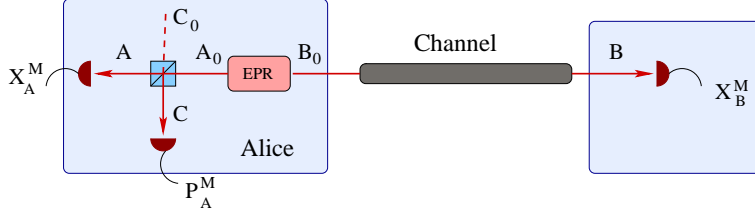


Figure 9.5: Entanglement based scheme of the protocol based on Alice sending coherent states and Bob applying homodyne detection. Alice generation of coherent states is replaced by an EPR state where Alice applies an heterodyne detection on one half of the EPR and the other half is sent to Bob.

Alice and Bob Mutual Information

Following the calculations of the previous chapter Alice and Bob mutual information reads,

$$S(x_A^M : x_B) = \frac{1}{2} \log \left[\frac{V_A + 1}{V_{A|B} + 1} \right] = \frac{1}{2} \log \left[\frac{V + \chi}{\chi + 1} \right], \quad (9.21)$$

which is the same for Direct and Reverse Reconciliation.

Eve Information: Direct Reconciliation

Eve quantum information on Alice measurements

$$S(x_A^M : E) = S(E) - S(E | x_A^M), \quad (9.22)$$

can be calculated in a similar way as before. The term $S(E) = S(AB)$ is exactly the same as in equation (9.11), while the evaluation of $S(E | x_A^M)$ is slightly more complex. After Alice projective measurement x_A the system BCE is pure, giving $S(E | x_A^M) = S(BC | x_A^M)$, where C is Alice auxiliary mode used at the heterodyne detection. $S(BC | x_A^M)$ being a function of the symplectic eigenvalues $\lambda_{3,4}$ of $\gamma_{BC}^{x_a}$. The covariance matrix $\gamma_{BC}^{x_a}$ being the result of applying homodyning over A after mixing A and C in a balanced beamsplitter:

$$\gamma_{ACB} = [S_{AC}^{BS} \otimes \mathbb{I}_B]^T \gamma_{A_0 C_0 B} [S_{AC}^{BS} \otimes \mathbb{I}_B] \quad (9.23)$$

where S_{AC}^{BS} is the symplectic transformation of the balanced beamsplitter. After Alice homodyne detection, the covariance matrix $\gamma_{BC}^{x_a}$ reads (using equation (2.57)),

$$\gamma_{BC}^{x_a} = \begin{pmatrix} b - c^2/a & 0 & \sqrt{2}c/(a+1) & 0 \\ 0 & b & 0 & -c/\sqrt{2} \\ \sqrt{2}c/(a+1) & 0 & 2a/(a+1) & 0 \\ 0 & -c/\sqrt{2} & 0 & (a+1)/2 \end{pmatrix}. \quad (9.24)$$

Then the conditional von Neumann entropy reads,

$$S(BC | x_A^M) = G[(\lambda_3 - 1)/2] + G[(\lambda_4 - 1)/2], \quad (9.25)$$

where

$$\lambda_{3,4}^2 = \frac{1}{2} \left[A \pm \sqrt{A^2 - 4B} \right], \quad (9.26)$$

where we have used the notation

$$\begin{aligned} A &= \frac{1}{a+1}[a + bD + \Delta] \\ B &= \frac{D}{a+1}[b + D], \end{aligned} \quad (9.27)$$

with D and Δ defined in equations (9.14).

Reverse Reconciliation

Bob still applying homodyne measurement, Eve quantum information on Bob measurement

$$S(x_B:E) = S(E) - S(E|x_B), \quad (9.28)$$

is then exactly the same as in the case of Reverse Reconciliation of the protocol based on squeezed states and homodyne measurement (9.3).

Squeezed States and Heterodyne Detection

The protocol based on squeezed states and heterodyne detection is equivalent to an entanglement based scheme where Alice applies an homodyne measurement over mode A and Bob applies heterodyne measurements over mode B , as shown in Fig. 9.6. The heterodyne measurement is modeled by combining mode B and a vacuum ancillae C in a balanced beamsplitter and measuring x on B and p on C . The secret key rate

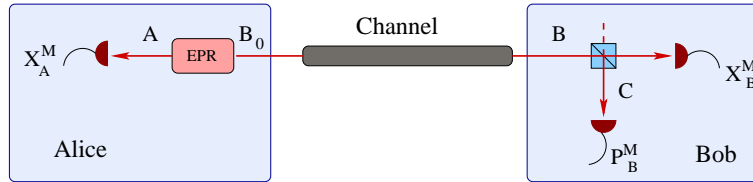


Figure 9.6: Entanglement based scheme of the protocol based on Alice sending squeezed states and Bob applying heterodyne detection. Alice generation of squeezed is replaced by an EPR state where Alice applies an homodyne detection on one half of the EPR and the other half is sent to Bob.

then reads,

$$K = S(x_A:x_B^M) - I_E, \quad (9.29)$$

where Eve information I_E reads $S(x_A:E)$ for DR and $S(x_B^M:E)$ for RR.

Alice and Bob Mutual Information

Following the calculations of the previous chapter Alice and Bob mutual information reads,

$$S(x_A:x_B^M) = \frac{1}{2} \log \left[\frac{V_B + 1}{V_{B|A} + 1} \right] = \frac{1}{2} \log \left[\frac{T(V + \chi) + 1}{T(\chi + 1/V) + 1} \right], \quad (9.30)$$

which is the same for Direct and Reverse Reconciliation.

Eve Information: Direct Reconciliation

Alice still applying homodyne measurements, Eve quantum information on Alice measurements

$$S(x_A:E) = S(E) - S(E|x_A), \quad (9.31)$$

is then exactly the same as in the case of Direct Reconciliation of squeezed states and homodyne measurement.

Eve Information: Reverse Reconciliation

Eve quantum information on Bob measurement

$$S(x_B^M:E) = S(E) - S(E|x_B^M), \quad (9.32)$$

can be calculated in a similar way as we did for Direct Reconciliation in the protocol based on coherent states and homodyning. By symmetry we have just to change $a \leftrightarrow b$ in $\gamma_{BC}^{x_a}$ to obtain $\gamma_{AC}^{x_b}$. It is then straightforward to calculate the conditional von Neumann entropy,

$$S(BC|X_B^M) = G[(\lambda_3 - 1)/2] + G[(\lambda_4 - 1)/2], \quad (9.33)$$

where

$$\lambda_{3,4}^2 = \frac{1}{2} \left[A \pm \sqrt{A^2 - 4B} \right], \quad (9.34)$$

where we have used the notation

$$\begin{aligned} A &= \frac{1}{b+1} [b + aD + \Delta], \\ B &= \frac{D}{b+1} [a + D]. \end{aligned} \quad (9.35)$$

Coherent States and Heterodyne Detection

The protocol based on coherent states and heterodyne detection [194] is equivalent to an entanglement based scheme where Alice applies an heterodyne measurement over mode A_0 and Bob applies heterodyne measurement over mode B' , as shown in Fig. 9.7. The secret key rate then reads,

$$K_{CE} = S(x_{A_0}^M, p_{A_0}^M : x_B^M, p_B^M) - I_E, \quad (9.36)$$

where Eve information I_E reads $S(x_{A_0}^M, p_{A_0}^M : E)$ for DR and $S(x_B^M, p_B^M : E)$ for RR.

Alice and Bob Mutual Information

Following the calculations of the previous chapter Alice and Bob mutual information reads,

$$S(x_{A_0}^M, p_{A_0}^M : x_B^M, p_B^M) = 2S(x_{A_0}^M : x_B^M) \quad (9.37)$$

$$= \log \left[\frac{V_B + 1}{V_{B|A^M} + 1} \right] = \log \left[\frac{T(V + \chi) + 1}{T(\chi + 1) + 1} \right], \quad (9.38)$$

which is the same for Direct and Reverse Reconciliation.

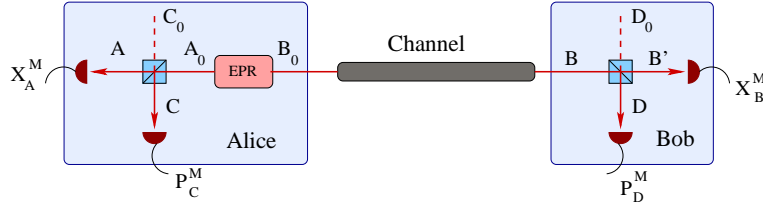


Figure 9.7: Entanglement based scheme of the protocol based on Alice sending coherent states and Bob applying heterodyne detection. Alice generation of coherent states is replaced by an EPR state where Alice applies an heterodyne detection on one half of the EPR and the other half is sent to Bob. Alice heterodyne measurement of mode A_0 is implemented by mixing A_0 with the auxiliary vacuum C_0 into a balanced beamsplitter and homodyning the outputs A and C to obtain x_A^M and p_C^M respectively. Similarly Bob heterodyne of mode B is implemented by mixing B with the auxiliary vacuum D_0 into a balanced beamsplitter and homodyning the outputs B' and D to obtain x_B^M and p_D^M respectively.

Eve Information: Direct Reconciliation

Eve quantum information on Alice measurement

$$S(x_{A_0}^M, p_{A_0}^M : E) = S(E) - S(E|x_{A_0}^M, p_{A_0}^M), \quad (9.39)$$

can be calculated in a similar way as before. The term $S(E) = S(AB)$ is exactly the same as in equation (9.11). After Alice heterodyne detection over mode A_0 , (homodyning A and C to obtain x_A^M and p_C^M , respectively) the system BE is pure. This gives $S(E|x_A^M, p_C^M) = S(B|x_A^M, p_C^M) = S(B|x_A, p_C)$, which is a function of the symplectic eigenvalue λ_3 of $\gamma_B^{x_a, p_a}$. The covariance matrix $\gamma_B^{x_a, p_a}$ results from applying a projective measurement over A (x_A) and C (p_C) over γ_{ACB} of equation (9.23), which using equation (2.57) reads,

$$\gamma_{BC}^{x_a} = (b - c^2/(a+1))\mathbb{I}. \quad (9.40)$$

Then the conditional von Neumann entropy reads,

$$S(B|x_A^M, p_C^M) = G[(\lambda_3 - 1)/2], \quad (9.41)$$

where

$$\lambda_3 = b - c^2/(a+1) = T(\chi + 1). \quad (9.42)$$

Eve Information: Reverse Reconciliation

Eve quantum information on Bob measurement

$$S(x_B^M, p_B^M : E) = S(E) - S(E|x_B^M, p_B^M), \quad (9.43)$$

is calculated in the same way as we did for Direct Reconciliation. After Bob heterodyne detection over modes B' (x_B) and D (p_D) the system AE is pure, giving $S(E|x_B^M, p_B^M) = S(E|x_{B'}, p_D) = S(A|x_{B'}, p_D)$, where D is Alice auxiliary mode. $S(A|x_{B'}, p_D)$ being the symmetric counterpart of $S(B|x_A, p_C)$ calculated for Direct Reconciliation, we just have to change $a \leftrightarrow b$ in $\gamma_B^{x_a, p_a}$ to obtain $\gamma_A^{x_b, p_b}$. Then the conditional von Neumann entropy reads,

$$S(A|x_{B'}, p_D) = G[(\lambda_3 - 1)/2], \quad (9.44)$$

where

$$\lambda_3 = a - c^2/(b+1) = \frac{T(V\chi + 1) + 1}{T(V + \chi) + 1}. \quad (9.45)$$

Comparison of the Different Protocols

For a lossy channel, (no excess noise, $\epsilon = 0$), we observe in Fig. 9.8 that the protocol

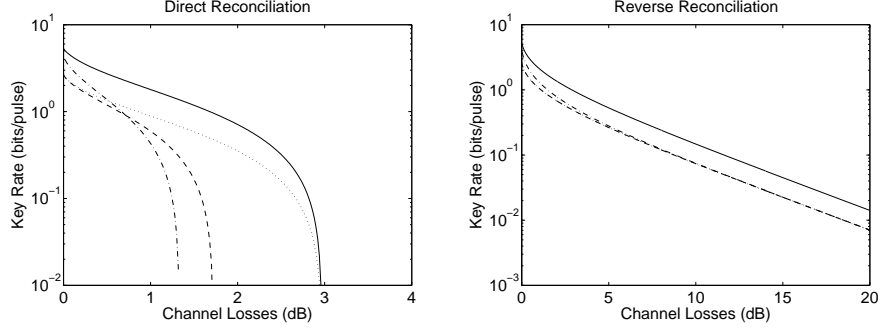


Figure 9.8: Secret key rate as a function of the channel losses (measured in dB) for a lossy channel ($\epsilon = 0$) for all the Gaussian protocols: squeezed states and homodyne detection (solid line), coherent states and homodyne detection (dotted line), coherent states and heterodyne detection (dot-dashed line) and squeezed states and heterodyne detection (dashed line). The curves are plotted for experimental realistic modulation $V = 40$.

based on squeezed states and homodyne detection [44] gives the highest rates for Direct and Reverse Reconciliation. Interestingly, compared to individual attacks where the no basis switching protocol [194] performed nearly as well as the protocols based on squeezed states and homodyning, here it performs much worse. Its performance, unless for low losses ($< 1\text{dB}$ in DR and $< 3\text{dB}$ in RR), is now comparable to that of coherent states and homodyning [96] in Reverse Reconciliation and that of squeezed states and heterodyne detection in Direct Reconciliation (which is much worse than coherent states and homodyning for DR). The protocol based on coherent states and

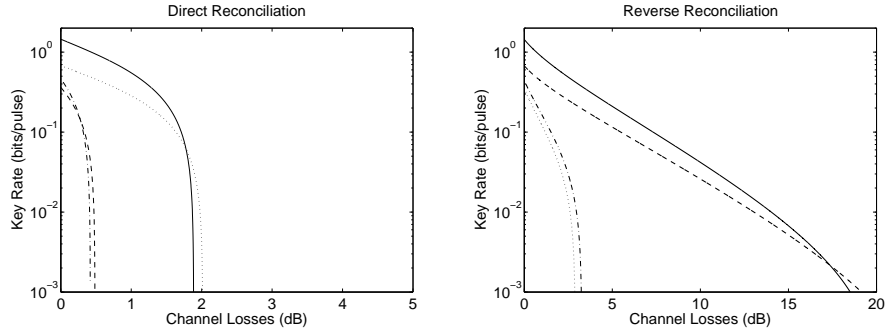


Figure 9.9: Secret key rate as a function of the channel losses (measured in dB) for a noisy channel ($\epsilon = 0.25$), for all the Gaussian protocols: squeezed states and homodyne detection (solid line), coherent states and homodyne detection (dotted line), coherent states and heterodyne detection (dot-dashed line) and squeezed states and heterodyne detection (dashed line). The curves are plotted for experimental realistic modulation $V = 40$.

homodyning and the protocols based on squeezed states and heterodyne being a noisy version of the protocol based on squeezed states and homodyne measurement, one would expect them to give worse secret key rates for any channel. Surprisingly that

is not the case as we can see in Fig. 9.9, where we observe that after some threshold coherent states and homodyne performs better in Direct Reconciliation and squeezed states and heterodyne in Reverse Reconciliation.

Noise Can Be Helpful

Looking carefully at the different steps of the Reverse Reconciliation protocol using squeezed states and heterodyne measurement, we observe that during the sifting phase Bob throws away one of his two measurements (either x_B^M or p_B^M) of the conjugate quadratures depending on the measurement done by Alice. This operation translates mathematically to tracing out the mode that does not correspond to the correct quadrature measurement (mode C on Fig.9.6). This implies that our protocol can be seen as a noisy version of the protocol based on squeezed states and homodyne detection where Bob applies a 50% lossy channel (not controlled by Eve) before his homodyne measurement. This is a clear demonstration that for reverse Reconciliation adding some noise on Bob side not controlled by Eve could be beneficial and increase the secret key rate, a phenomenon already known for discrete variables (see [158]) but unknown for continuous variables. Using the entanglement-based description it is trivial to see that this phenomena has a symmetric counterpart in Direct Reconciliation, where the protocol based on coherent states and homodyne detection can be seen as the squeezed state and homodyne protocol with a 50% lossy measurement on Alice side, outperforming it for sufficiently high excess noise and losses.

Remark that in order to be beneficial, the noise must be added on the reference partner of the reconciliation, Alice in DR and Bob in DR. Otherwise, the noise only affects Alice and Bob mutual information without decreasing Eve information on the final key. This explains why coherent states and homodyne (squeezed states and heterodyne) performs worse than squeezed states and homodyne in RR (DR).

Tolerable Excess Noise

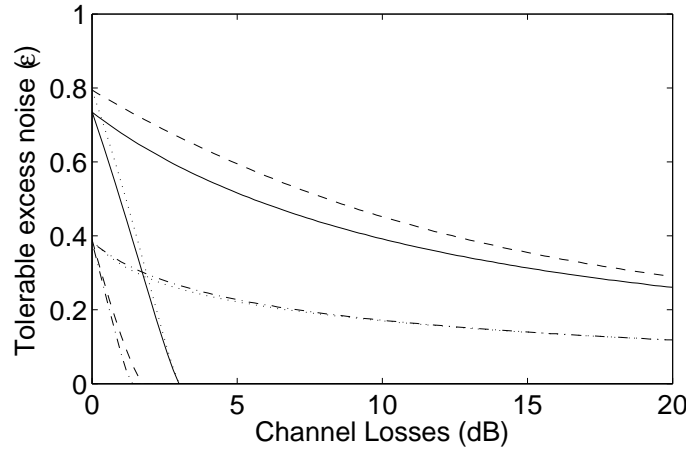


Figure 9.10: Tolerable excess noise ϵ as a function of the channel losses (measured in dB) for high modulation ($V = 1000$) for all the Gaussian protocols: squeezed states and homodyne detection (solid line), coherent states and homodyne detection (dotted line), coherent states and heterodyne detection (dot-dashed line) and squeezed states and heterodyne detection (dashed line). The curves vanishing at (or above) 3dB correspond to DR, whereas the rest refers to RR.

In Fig. 9.10 we plot for all the proposed protocols the resistance to noise as a function of the losses, which gives the excess noise ϵ given a null secret key K for a given channel of transmittance T . As for individual attacks we remark that all the Direct Reconciliation protocols have a maximal range of 3 dB where for Reverse Reconciliation there is no theoretical limitation to the range. Having in mind the results of the preceding subsection, it is not a surprise that the protocol based in squeezed states and heterodyne detection (coherent states and homodyne detection) gives the optimal resistance to noise in Reverse Reconciliation (Direct Reconciliation), as shown in Fig. 9.10.

In Fig. 9.11 we compare the two protocols based on homodyne detection on Bob side and the two protocols based on squeezed states in RR, the dotted line giving the threshold where the respective noisy version outperforms the protocol based on squeezed states and homodyne detection.

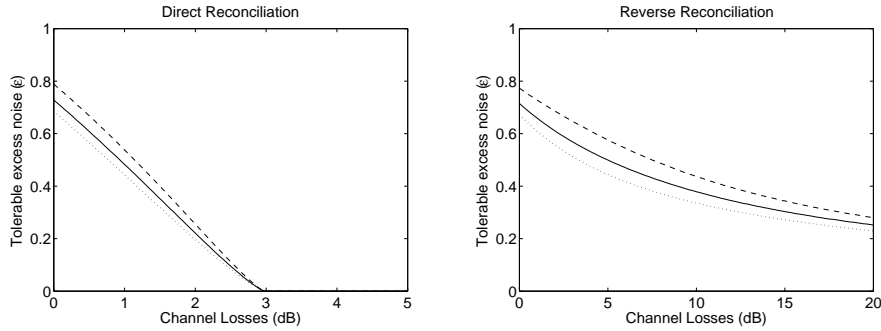


Figure 9.11: Tolerable excess noise ϵ as a function of the channel losses (measured in dB) for high modulation ($V = 40$). For DR we compare the protocol based on squeezed states and homodyne detection (solid line) and coherent states with homodyne detection (dashed line), where for RR we compare the protocol based on squeezed states and homodyne detection (solid line) and squeezed states with heterodyne detection (dashed line). The dotted line shows the channels (defined by the parameters T, ϵ) giving the same secret key rate for both protocols.

9.4 Fighting Noise with Noise

In a recent paper [158] the authors pointed a surprising effect not previously observed. Studying the protocols BB84 [20], BB92 [19] and the six-state protocol [38, 14] they realized that Alice could increase the performance of the protocols by adding some noise to her data before the error correction processing. Strikingly this additional classical noise makes the protocol more robust against noise in the quantum channel. More precisely the authors showed that for each quantum channel there is an optimal classical added noise that Alice has to add in order to optimize the secret key rate. It is easy to understand that the operation of addition of classical noise must be done by the partner who will be the reference during the error correction processing, as adding noise on the other partner only decreases the mutual information of the trustful parties without having any effect on Eve. In a very recent paper [154] the authors gave an explanation to this phenomenon by generalizing the Shor-Preskill proof of unconditional security of BB84 [180] that combines an entanglement based description of the protocol and entanglement distillation with CSS codes [136]. The authors exploit the result of [107] that shows that entanglement distillation is not a necessary condition for security of QKD, as it is enough to distill a more general class of states called private states. The

authors show in [154] that the entanglement based description of the states obtained after the addition of noise on Alice side can be distilled to a class of private states, increasing the secret key rate. The analysis done in the preceding section for existing Gaussian protocols shows that this striking effect exists also in continuous variable QKD and explains some differences between the existing protocols that were not understood previously. In the following we are going to generalize the previous protocol to a general Gaussian phase-insensitive noise added by either Alice (DR) or Bob (RR) in order to optimize the secret key rate.

Reverse Reconciliation

In Reverse Reconciliation the optimal protocol would be a source of squeezed states on Alice side combined with an inefficient homodyne measurement on Bob side, where the efficiency of the detection and the electronic noise are chosen in order to optimize the secret key rate. Fig. 9.12 shows an entanglement-based description of this new protocol, where the efficiency of the detection is modeled by a beamsplitter of transmittance T_B and the electronic noise v is modeled by a thermal noise (variance N_B) added at the second input of the beamsplitter T_B ($v = (1 - T_B)(N_B - 1)$). In the following we will

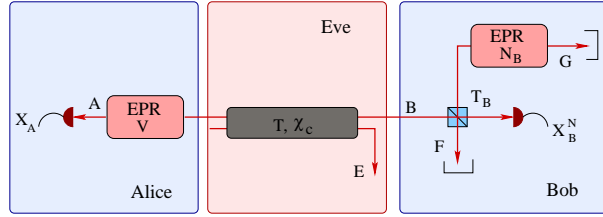


Figure 9.12: Generalized entanglement-based description of the protocol with general Gaussian added noise on Bob side. The source of squeezed states on Alice side is replaced by an entangled pair (EPR) of squeezing parameter V followed by an homodyne measurement by Alice on half of the pair. The other half of the EPR is sent to Bob through the channel of transmittance T and added noise $\chi = (1 - T)/T + \epsilon$. Before Bob homodyne detection the state received by Bob is mixed with a thermal state (mode F) of variance N on a beamsplitter of transmittance T_B . The additional noise referred to the input of the detector is $\chi_{DB} = (1 - T_B)N_B/T_B$.

show that the effect of the Gaussian noisy measurement (T_B, v) on the key rate depends only on one parameter, the added noise referred to the input of the measurement device ($\chi_{DB} = (1 + v)/T_B - 1$). Therefore all the combinations of the parameters (T_B, v) giving the same χ_{DB} are equivalent in terms of their effect on the secret key rate. An alternative implementation would be a perfect homodyne detection followed by a classical Gaussian noise of variance χ_{DB} added on Bob data. The secret key rate reads

$$K = S(x_A : x_B^N) - S(x_B^N : E). \quad (9.46)$$

Note that x_B^N stands for the noisy measurement of the x quadrature of mode B .

Alice and Bob Mutual Information

Alice and Bob mutual information, generalizing the results of chapter 8 reads,

$$I(x_A : x_B^N) = \frac{1}{2} \log \left[\frac{V_B^N}{V_{B^N|A}} \right] = \frac{1}{2} \log \left[\frac{T_B V_B + (1 - T_B) N_B}{T_B V_{B|A} + (1 - T_B) N_B} \right] \quad (9.47)$$

$$= \frac{1}{2} \log \left[\frac{V + \chi_T}{\chi_T + 1/V} \right]. \quad (9.48)$$

where $\chi_T = \chi + \chi_{D_B}/T$ is the total added noise referred to the input, χ the channel added noise and $\chi_{D_B} = (1 - T_B)N/T_B$ the detector added noise.

Eve Information

Eve quantum information on Bob measurement

$$S(x_B^N : E) = S(E) - S(E | x_B^N), \quad (9.49)$$

can be easily calculated generalizing the technique used previously. The term $S(E) = S(AB)$ is exactly the same as in equation (9.11). Secondly, after Bob projective measurement x_B the system $AEFG$ is pure, then $S(E | x_B^N) = S(AFG | x_B^N)$. In order to calculate $S(AFG | x_B^N)$ we need the covariance matrix $\gamma_{AFG}^{x_b}$, which results from applying a projective measurement on mode B on γ_{ABFG} ,

$$\gamma_{ABFG} = [\mathbb{I}_A \otimes S_{AC}^{BS} \otimes \mathbb{I}_G]^T \gamma_{AB} \otimes \gamma_{FG} [\mathbb{I}_A \otimes S_{AC}^{BS} \otimes \mathbb{I}_G], \quad (9.50)$$

where γ_{AB} was given in Eq. (9.5) and γ_{FG} is the covariance matrix of an EPR state of variance $N = T_B v / (1 - T_B)$. The covariance matrix $\gamma_{AFG}^{x_b}$ reads,

$$\gamma_{AFG}^{x_b} = \begin{pmatrix} \gamma_A & \sigma_{AF} & \sigma_{AG} \\ \sigma_{AF}^T & \gamma_F & \sigma_{FG} \\ \sigma_{AG}^T & \sigma_{FG}^T & \gamma_G \end{pmatrix}, \quad (9.51)$$

where

$$\gamma_A = \begin{pmatrix} a - \frac{T_B c^2}{T_B b + (1 - T_B)N} & 0 \\ 0 & a \end{pmatrix}, \quad (9.52)$$

$$\gamma_F = \begin{pmatrix} \frac{bN}{T_B b + (1 - T_B)N} & 0 \\ 0 & (1 - T_B)b + T_B N \end{pmatrix}, \quad (9.53)$$

$$\gamma_G = \begin{pmatrix} N - \frac{(1 - T_B)(N^2 - 1)}{T_B b + (1 - T_B)N} & 0 \\ 0 & N \end{pmatrix}, \quad (9.54)$$

$$\sigma_{AF} = \begin{pmatrix} \frac{\sqrt{1 - T_B} N c}{T_B b + (1 - T_B)N} & 0 \\ 0 & -\sqrt{1 - T_B} c \end{pmatrix}, \quad (9.55)$$

$$\sigma_{AG} = \begin{pmatrix} \frac{\sqrt{T_B(1 - T_B)(N^2 - 1)}b}{T_B b + (1 - T_B)N} & 0 \\ 0 & 0 \end{pmatrix}, \quad (9.56)$$

and

$$\sigma_{FG} = \begin{pmatrix} \frac{\sqrt{T_B(N^2 - 1)}b}{T_B b + (1 - T_B)N} & 0 \\ 0 & -\sqrt{T_B(N^2 - 1)} \end{pmatrix}. \quad (9.57)$$

$S(AFG|x_B^N)$ being a function of the symplectic eigenvalues $\lambda_{3,4,5}$, it can be calculated from the symplectic invariants Δ_1^3, Δ_2^3 and Δ_3^3 defined in Chapter 2. After some calculation we obtain,

$$\Delta_1^3 = \frac{1}{b + \chi_{D_B}} [2b + aD + \chi_{D_B}(\Delta + 1)], \quad (9.58)$$

$$\Delta_2^3 = \frac{1}{b + \chi_{D_B}} [b + 2aD + \chi_{D_B}(D^2 + \Delta)], \quad (9.59)$$

$$\Delta_3^3 = \frac{D}{b + \chi_{D_B}} [a + \chi_{D_B}D]. \quad (9.60)$$

Interestingly the symplectic invariants satisfy the relation $1 - \Delta_1^3 + \Delta_2^3 - \Delta_3^3 = 0$ giving $\lambda_5 = 1$ ($G(\lambda_5) = 0$), as the symplectic eigenvalues are solution of the polynomial $z^3 - \Delta_1^3 z^2 + \Delta_2^3 z - \Delta_3^3 = 0$. Then, $\lambda_{3,4}^2$ are solutions of the second order polynomial $z^2 - Az + B = 0$, where

$$\begin{aligned} A &= \Delta_1^3 - 1 = \frac{1}{b + \chi_{D_B}} [b + aD + \chi_{D_B}\Delta], \\ B &= \Delta_3^3 = \frac{D}{b + \chi_{D_B}} [a + \chi_{D_B}D]. \end{aligned} \quad (9.61)$$

We observe that the dependence of A and B on the Gaussian noisy measurement implemented by Bob only depends on the parameter χ_{D_B} , as we mentioned previously. There are then many combinations of the parameters (T_B, v) that give the same result. As expected, Bob heterodyne detection ($T_B = 1/2, N = 0$) corresponds to $\chi_D = 1$ where equation (9.61) recovers the solution of equation (9.35) for the protocol based on squeezed states and heterodyne detection.

Results

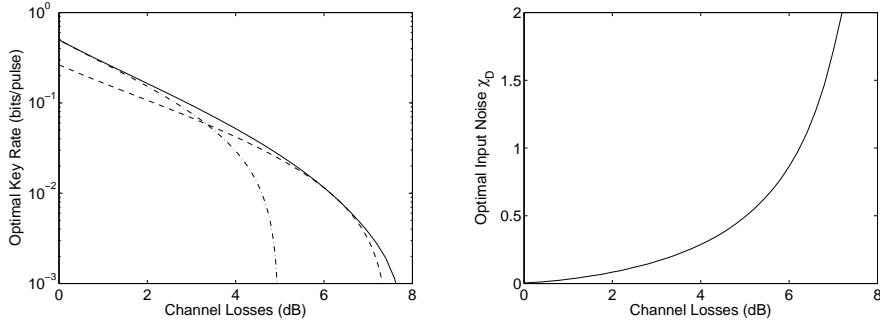


Figure 9.13: a) Secret key rate as a function of the channel losses (measured in dB) for a channel excess noise ($\epsilon = 0.5$) for the Gaussian RR protocols: squeezed states and homodyne detection (dot-dashed line), squeezed states and heterodyne detection (dashed line) and the optimization over Bob's measurement added noise χ_{D_B} (solid line). b) Optimal choice of χ_{D_B} that maximize the secret key rate at a). All the curves have been plotted for experimental realistic modulation $V = 40$.

By correctly tuning Bob added noise χ_{D_B} it is possible to optimize the secret key rate as we show in Fig. 9.13, where we compare the previous protocols based on Alice generating squeezed states with the optimization over the noise on Bob side.

In Fig. 9.14 we compare the maximal tolerance to noise as a function of the losses of

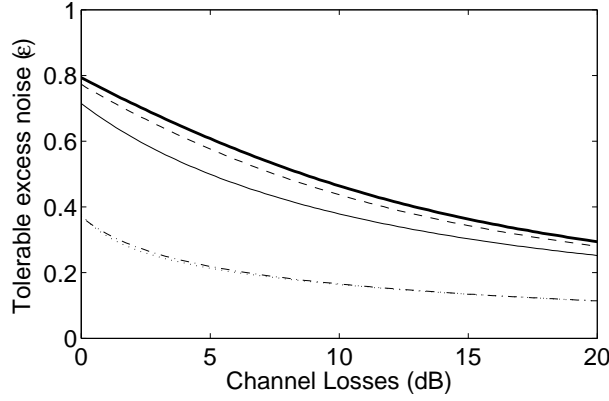


Figure 9.14: Tolerable excess noise ϵ as a function of the channel losses (measured in dB) for modulation ($V = 40$) for all the Gaussian RR protocols: optimization over Bob's measurement added noise χ_{DB} (bold solid line), squeezed states and homodyne detection (solid line), coherent states and homodyne detection (dotted line), coherent states and heterodyne detection (dot-dashed line) and squeezed states and heterodyne detection (dashed line).

the channel for the previously existing Gaussian protocols and the optimized protocol. We observe that the optimization slightly improves the resistance of the protocol based on squeezed states and heterodyne detection.

Direct Reconciliation

Now we are going to consider the Direct Reconciliation counterpart of the previous effect. Fig. 9.15 shows an entanglement-based description of this new protocol where Alice replaces his homodyne measurement over half of the EPR pair by an inefficient homodyne measurement of efficiency T_A and added noise χ_{DA} . The secret key rate

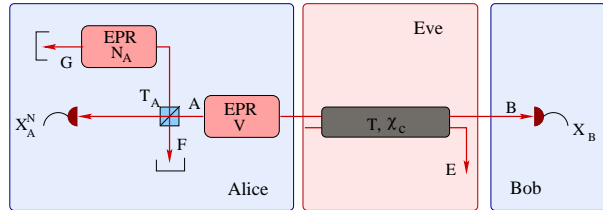


Figure 9.15: Generalized entanglement-based description of the protocol with general Gaussian added noise on Alice side. The source of squeezed states on Alice side is replaced by an entangled pair (EPR) of squeezing parameter V followed by an inefficient homodyne measurement (T_A, χ_{DA}) by Alice on half of the pair. The other half of the EPR is sent to Bob through the channel. The additional noise referred to the input of the detector reads $\chi_{DA} = (1 - T_A)N/T_A$.

then reads,

$$K = S(x_A^N : x_B) - S(x_A^N : E). \quad (9.62)$$

Alice and Bob Mutual Information

Alice and Bob mutual information, generalizing the results of chapter 8 reads,

$$I(x_A^N : x_B) = \frac{1}{2} \log \left[\frac{V_A^N}{V_{A^N|B}} \right] = \frac{1}{2} \log \left[\frac{T_A V_A + (1 - T_A) N_A}{T_A V_{A|B} + (1 - T_A) N_A} \right] \quad (9.63)$$

$$= \frac{1}{2} \log \left[\frac{(V + \chi)(V + \chi_{D_A})}{V(\chi + \chi_{D_A} + \chi\chi_{D_A} + 1)} \right], \quad (9.64)$$

where χ is the channel added noise and $\chi_{D_A} = (1 - T_A)N_A/T_A$ Alice detector added noise.

Eve Information

Eve quantum information on Bob measurement reads,

$$S(x_A^N : E) = S(E) - S(E|x_A^N). \quad (9.65)$$

It can be easily calculated generalizing the technique used previously. The term $S(E) = S(AB)$ is exactly the same as in equation (9.11). Secondly, after Alice projective measurement x_A^N the system $BEFG$ is pure, then $S(E|x_A^N) = S(BFG|x_A^N)$. $S(BFG|x_A^N)$ being the symmetric counterpart of $S(AFG|x_B^N)$ calculated in the Reverse Reconciliation case, we just have to change $a \leftrightarrow b$ in equation (9.61). The square of the symplectic eigenvalues $\lambda_{3,4}^2$ being solutions of the second order polynomial $z^2 - Az + B = 0$, where

$$\begin{aligned} A &= \frac{1}{a + \chi_{D_A}} [a + bD + \chi_{D_A} \Delta], \\ B &= \frac{D}{a + \chi_{D_A}} [b + \chi_{D_A} D]. \end{aligned} \quad (9.66)$$

Notice that Alice heterodyne detection corresponds with $\chi_{D_A} = 1$ recovering the solution of equation (9.27), as expected.

Results

By correctly tuning Alice added noise χ_{D_A} it is possible to optimize the secret key rate as we show in Fig. 9.16. where we compare the previous protocols based on Bob applying homodyne measurement with the optimization over Alice measurement with phase insensitive noise, which is equivalent to a noisy prepare-and-measure scheme.

In Fig. 9.17 we compare the maximal tolerance to noise as a function of the losses of the channel for the previously existing Gaussian protocols and the optimized protocol. We observe that the optimization slightly improves the resistance of the protocol based on coherent states and homodyne detection and does not succeed to pass the 3dB limitation to the range in Direct Reconciliation protocols, as expected.

Alice Source

In the case of Reverse Reconciliation the translation from the entanglement based to the prepare-and-measure scheme was trivial, Bob had either to apply a noisy measurement of added noise χ_{D_B} or a perfect homodyne detection followed by a classical noise of variance χ_{D_B} . In the case of Direct Reconciliation Alice can obviously send squeezed states and apply a classical added noise of variance χ_{D_A} over her data. This solution has two disadvantages, we need a source of squeezed states and more random numbers to implement the classical noise.

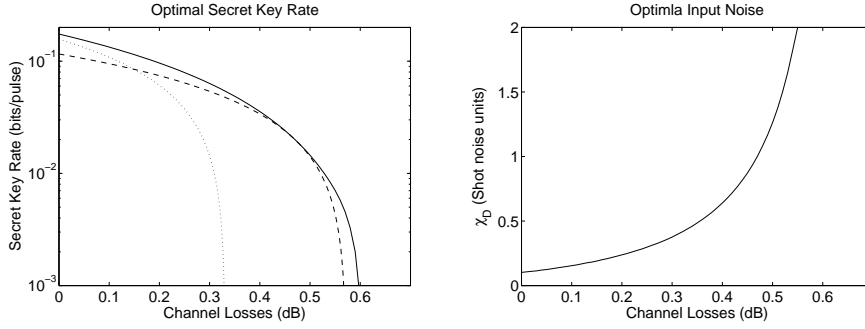


Figure 9.16: a) Secret key rate as a function of the channel losses (measured in dB) for a channel excess noise ($\epsilon = 0.65$) for the Gaussian DR protocols: squeezed states and homodyne detection (dot-dashed line), coherent states and homodyne detection (dashed line) and the optimization over Alice's measurement added noise χ_{DA} (solid line). b) Optimal choice of χ_{DA} that maximizes the secret key rate on a). The curves have been plotted for experimental realistic modulation $V = 40$.

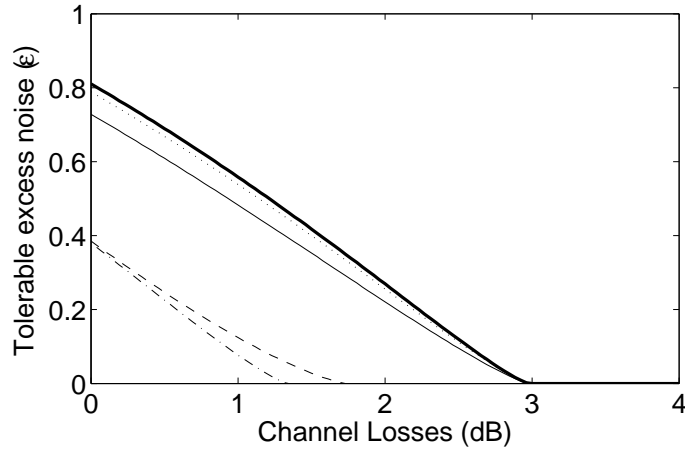


Figure 9.17: Tolerable excess noise ϵ as a function of the channel losses (measured in dB) for modulation $V = 40$ for all the Gaussian DR protocols: optimization over Alice's measurement added noise χ_{DA} (bold solid line), squeezed states and homodyne detection (solid line), coherent states and homodyne detection (dotted line), coherent states and heterodyne detection (dot-dashed line) and squeezed states and heterodyne detection (dashed line).

Interestingly for $\chi_{DA} \geq 1$ there is an easier implementation based on noisy coherent states. The noise $\chi_{DA} = 1$ can be obtained by fixing $T_A = 1/2$ which is equivalent to a source of coherent states. Then it's possible to cover all the range $\chi_{DA} \geq 1$ by just sending noisy coherent states (displaced thermal states), where the variance of the added noise is a function of χ_{DA} . Unfortunately for $\chi_{DA} \leq 1$ we cannot avoid using a source of squeezed states.

9.5 Fiber Optic Implementation

The joint collaboration of the QuIC center of the ULB with the Charles Fabry laboratory of the Institut d'Optique d'Orsay, Thalès Research & Technologies, and GeorgiaTech-Metz, has recently successfully implemented a fully functional CV-QKD system ready to field implementation, generating secret keys at a rate of more than 2 kb/s over 25 km of optical fiber [128]. This CV-QKD system is based on a source of coherent states at Alice side combined with an homodyne detection at Bob side, using a reverse reconciliation protocol. Using a source of coherent states has the advantage over squeezed states that we can use standard (low-cost) telecom optical components, where the measurement on Bob side is limited to homodyning in order to simplify the implementation. Finally we choose a reverse reconciliation protocol in order to beat the 3dB limit imposed to reverse reconciliation protocols.

The security analysis of this CV-QKD system is strikingly similar to the generalized protocol of the preceding section 9.4 based on a source of squeezed states and a general phase insensitive noise on Bob side. Alice and Bob mutual information reads,

$$I(x_A^M : x_B^N) = \frac{1}{2} \log \left[\frac{V_B^N}{V_{B^N|A^M}} \right] = \frac{1}{2} \log \left[\frac{T_B V_B + (1 - T_B) N_B}{T_B V_{B|A^M} + (1 - T_B) N_B} \right] \quad (9.67)$$

$$= \frac{1}{2} \log \left[\frac{V + \chi_T}{\chi_T + 1} \right]. \quad (9.68)$$

where $\chi_T = \chi + \chi_{D_B}/T$ is the total added noise referred to the input, χ the channel added noise and $\chi_{D_B} = (1 - T_B)N/T_B$ the detector added noise. Eve information on Bob measurement is given by equation (9.49), as Alice changing from squeezed states to coherent states does not change Eve information on Bob.

Part III

Conclusion and Perspectives

Conclusion and Perspectives

In this thesis we have studied different aspects of the novel field of quantum information with continuous variables. The higher efficiency and bandwidth of homodyne detection combined with the easiness of the generation and manipulation of Gaussian states makes continuous-variable quantum information a promising and flourishing field of research. This dissertation is divided in two parts. The first part explores some applications of the *photon subtraction* operation, while the second develops a detailed analysis of an important family of *continuous-variable quantum key distribution* protocols, namely those based on Gaussian modulation of Gaussian states.

Photon subtraction After a detailed introduction to quantum optics and phase-space representation of continuous-variable systems we have presented two different applications of the *photon subtraction* operation, being to date one of the simplest techniques to generate non-Gaussian states of light.

In Chapter 4 we have shown that an arbitrary single-mode state of light can be engineered starting from a squeezed vacuum state and applying a sequence of displacements and single-photon subtractions, followed by a final squeezing operation. The recent demonstration of single-photon subtraction from a single-mode squeezed vacuum [142, 197] provides a strong evidence of the practical feasibility of our scheme, which will be much easier than previous proposals as it does not require single-photon sources and can operate with low-efficiency photodetectors. Generalizing this technique to two or more modes of light deserves further investigation, as it could lead to simpler way of generating highly entangled states of light, which could improve the existing proposal of loophole-free Bell test using homodyne detection. Two recent experiments [142, 141] show that photon subtraction is a feasible operation with a vast range of promising applications such as "Schrödinger cat" generation, entanglement distillation and implementation of Bell tests.

In Chapter 5 we proposed an experimentally feasible setup allowing for a loophole-free Bell test with efficient homodyne detection. This proposal is probably the simplest loophole-free Bell test experiment proposed so far based on quantum states of light and homodyne detection. We showed that a violation of Bell inequalities becomes possible using a non-Gaussian entangled state generated from a two-mode squeezed vacuum state by subtracting a single photon from each mode. We made a full analytical description of a realistic setup, studying in detail the influence of the detector inefficiencies, the electronic noise of homodyne detector and the efficiency of the mode filtering that must precede the photon subtraction. However, the class of schemes that we have studied is still somewhat restricted. The search of new non-Gaussian states easier to generate and/or giving better violations deserves future investigation. Even more interesting would be to find new Bell inequalities, using more and/or different measurements and/or different binning giving a higher violation, experimentally accessible with nowadays technology. We hope that the combination of improvements on the experimental side together with new theoretical ideas will lead to the first loophole-free Bell test using the continuous-variable paradigm in the near future.

Continuous-variable quantum key distribution Quantum key distribution (QKD) is the most promising and developed application of the novel field of quantum information. Over the past few years, an important research effort has been devoted to continuous-variable quantum key distribution protocols (CV-QKD), motivated by the prospects of realizing high-rate cryptosystems relying on homodyne detection instead of photon counting. Since the discovery of a protocol based on coherent states [96] these systems also have the advantage that they are based on standard (low-cost) telecom optical components, using a simple laser as a source is enough to distribute a secret key. The joint collaboration of the QuIC center of the ULB with the Charles Fabry laboratory of the Institut d'Optique d'Orsay, Thales Research & Technologies, and GeorgiaTech-Metz, has recently successfully implemented a fully functional CV-QKD system ready to field implementation, generating secret keys at a rate of more than 2 kb/s over 25 km of optical fiber [128]. This field implementation will be one of the four platforms of the first European Network for Secure Communication based on Quantum Cryptography, under development by the European consortium SECOQC.

Our contribution to this project was to prove the unconditional security of the family of Gaussian CV-QKD protocols ², giving at the same time a systematic way of calculating the secret key rates. This was achieved using a unified way of representing all the existing QKD protocols based on Gaussian modulation of squeezed (coherent) states by Alice and homodyne (heterodyne) detection by Bob for the two versions of one-way reconciliation (Direct Reconciliation and Reverse Reconciliation). Subsequently we showed that adding noise on the reference partner of the error correction post-processing can increase the secret key rate of a continuous-variable one-way QKD protocols, clarifying the pros and cons of the different protocols of the family of Gaussian protocols. Interestingly, for every quantum channel there is an optimal noise that must be added in order to optimize the secret key rate. Finally, we have completed the study of individual attacks, a weaker family of attacks compared to collective, that had remained open since the proposal of the no basis switching protocol [194] (with Alice sending coherent states and Bob performing heterodyne measurements). We have found that, in contrast with all other Gaussian protocols that had been studied so far, no individual attack exists that attains the security bounds deduced from the usual Heisenberg uncertainty relations, making these bounds unreachable in the present case. A tight bound was derived, both in direct and reverse reconciliation, and several explicit optical schemes that attain this bound have been exhibited.

The disadvantage of continuous variables compared to qubit-based QKD is the limited range of the existing protocols. The origin of this limitation is the sensitivity of the error correction post-processing to the unavoidable vacuum noise. Qubit-based QKD does not suffer from this effect because the vacuum noise is filtered out by the avalanche photodetectors, as only the events with a detected photon are used to generate the secret key. Unfortunately homodyne detection is not capable of such a filtering, generating useless data that only contribute to increase the noise, making the practical reconciliation more difficult. To overcome this problem, different protocols have been proposed recently [98, 123, 181], where a filtering stage is added after the homodyne detection in order to get rid of the harmful vacuum noise. Unfortunately, the security of those protocols is not fully understood, since it is based on the unproven assumption that the optimal attack is Gaussian. The ideal case would be to devise a non-Gaussian protocol implementing some filtering of the vacuum noise and keeping at the same time the tractability of Gaussian states calculations. This would be similar to what we did in the photon subtraction calculations, where we succeeded to model non-Gaussian states keeping the easiness of the calculations with Gaussian states.

²It is known for qubit-based QKD that security against collective attacks is enough to have unconditional security (see Chapter 8). It then seems very reasonable to conjecture that the same holds for CV-QKD.

Part IV

Appendices

Appendix A

The Church of the Larger Hilbert Space

The phrase "Going to the Church of the Larger Hilbert Space", coined by John Smolin, for the dilation construction of channels and states, is an extremely useful tool in quantum information theory. It is based on two key ideas; Firstly, every mixed state can be seen as being the partial trace of a pure state defined in a larger Hilbert space. Secondly, every quantum operation over a given quantum state can be described by a reversible interaction between the system and its environment. Finally, following Everett, the measurement can be considered as a type of quantum operation between the object and a pointer system.

Schmidt Decomposition

Any bipartite pure state $|\psi\rangle_{AB}$ can be written as,

$$|\psi\rangle_{AB} = \sum_i \lambda_i |i\rangle_A |i\rangle_B, \quad (\text{A.1})$$

where $|i\rangle_{A(B)}$ is an orthonormal state of systems $A(B)$ and λ_i are non-negative real numbers satisfying $\sum_i \lambda_i^2 = 1$ known as Schmidt coefficients.

Proof $|\psi\rangle_{AB}$ can be written

$$|\psi\rangle_{AB} = \sum_{j,k} a_{j,k} |j\rangle_A |k\rangle_B, \quad (\text{A.2})$$

using the singular value decomposition $a = vdw$, where v and w are unitary matrices and d is diagonal, $|i\rangle_A = \sum_j v_{ji} |j\rangle_A$ and $|i\rangle_B = \sum_k w_{ki} |k\rangle_B$ gives the result. \square

This result is extremely useful. For example, we see that $\rho_A = \sum_i \lambda_i^2 |i\rangle\langle i|_A$ and $\rho_B = \sum_i \lambda_i^2 |i\rangle\langle i|_B$ have exactly the same eigenvalues. Many important properties of quantum states, such as the entropy, are completely determined by the eigenvalues, so for pure bipartite systems it will be the same for both subsystems.

Purification

An extremely useful technique for quantum information calculations is the purification. If we are given a quantum state $\rho_A = \sum_i \lambda_i |i\rangle\langle i|_A$ of a quantum system A , it is possible

to introduce another system R (the reference) and define a pure state

$$|\psi\rangle_{RA} = \sum_i \sqrt{\lambda_i} |i\rangle_R |i\rangle_A, \quad (\text{A.3})$$

such that $\text{Tr}_R[|\psi\rangle\langle\psi|_{RA}] = \rho_A$. The idea is to define a pure state whose Schmidt coefficients are the square root of the eigenvalues of ρ_A .

Freedom of Purification Let $|\psi_1\rangle_{RA}$ and $|\psi_2\rangle_{RA}$ be two different purifications of ρ_A . It is easy to prove that there is a unitary operation U_R such that

$$|\psi_2\rangle_{RA} = U_R \otimes \mathbb{I}_A |\psi_1\rangle_{RA}, \quad (\text{A.4})$$

as both purifications have the same Schmidt decomposition up to a different orthonormal basis in system R .

Quantum Operations

A natural way to describe the dynamics of an open quantum system A is to regard it as arising from a unitary (reversible) interaction U_{AE} between the system A and the environment E which together form a closed system, as shown in Fig. A.1.

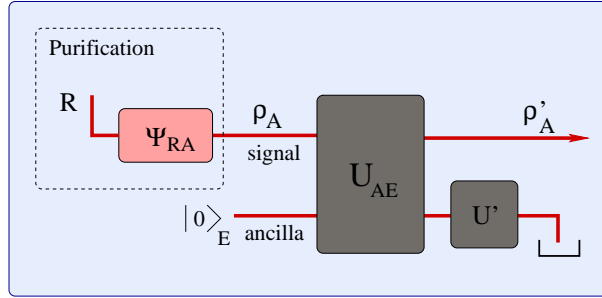


Figure A.1: Any quantum operation \mathcal{E} can be expressed in a Kraus operator representation $\{E_i\}$, which can be thought of as arising from a unitary evolution U_{AE} of the target quantum state ρ_A with an auxiliary system E and subsequently tracing E . The state ρ_A can be seen as the partial trace of its purification $|\psi\rangle_{AB}$. The Kraus representation $\{E_i\}$ of the quantum operation \mathcal{E} is unique up to a unitary operation (U') on the auxiliary system E .

The resulting state after applying the quantum operation, also called Completely Positive map (CP map), over the state ρ_A reads

$$\mathcal{E}(\rho) = \text{Tr}_E[U_{AE}(\rho_A \otimes |0\rangle\langle 0|_E)U_{AE}^\dagger], \quad (\text{A.5})$$

where there is no loss of generality in assuming that the environment starts in a pure state. Defining $|e_l\rangle$ as an orthonormal basis of system E we can rewrite the quantum operation

$$\mathcal{E}(\rho) = \sum_l \langle e_l | U_{AE} [\rho_A \otimes |0\rangle\langle 0|_E] U_{AE}^\dagger | e_l \rangle_E \quad (\text{A.6})$$

$$= \sum_l E_l \rho E_l^\dagger, \quad (\text{A.7})$$

where $E_l = \langle e_l | U_{AE} | e_0 \rangle_E$ is an operator on the state space of system A . The ensemble $\{E_l\}$ is called the Kraus operator representation of \mathcal{E} . The Kraus operators satisfy the so-called *completeness relation*,

$$\sum_l E_l^\dagger E_l = \mathbb{I}, \quad (\text{A.8})$$

which arises from the requirement $\text{Tr}[\mathcal{E}(\rho)] = 1$.

Unitary Representation Given a Kraus operator representation $\{E_l\}$, it is always possible to construct a corresponding unitary representation,

$$U_{AE} |\varphi\rangle_A \otimes |0\rangle_E = \sum_l E_l |\varphi\rangle_A \otimes |l\rangle_E, \quad (\text{A.9})$$

which preserves the inner product, therefore being unitary.

How many Kraus operators? All quantum operations \mathcal{E}_A defined on a Hilbert space of dimension d can be generated by a Kraus operator representation containing at most d^2 elements. In order to prove it, the key ingredient is to use the purification of system A , $|\psi\rangle_{RA}$. After applying the operation $\mathbb{I}_R \otimes \mathcal{E}_A$ to the initial state of system RA we obtain a mixed state ρ_{RA} which can be expanded as

$$\rho_{RA} = \sum_l \lambda_l |\varphi_l\rangle \langle \varphi_l|. \quad (\text{A.10})$$

One can show that each state $|\varphi_l\rangle$ is indeed associated with a Kraus operator E_l . Since ρ_{RA} has at most rank d^2 , \mathcal{E}_A has always a Kraus operator representation with at most d^2 elements. Therefore an ancillary system E in the unitary representation of dimension d^2 is enough. One can generalize the previous result to quantum operations \mathcal{E}_A with input and output Hilbert space of different dimensions; let fix us d and d' , it is then easy to prove that an ancillary system E of dimension dd' is enough.

How ambiguous? Suppose $\{E_i\}$ and $\{F_i\}$ are Kraus operators giving the quantum operations \mathcal{E}_A and \mathcal{F}_A . Then $\mathcal{E}_A = \mathcal{F}_A$ if and only if there is a unitary matrix U such that $E_i = \sum_j U_{ij} F_j$.

Consider the purification $|\psi\rangle_{RA}$ of the initial state ρ_A . After applying a quantum operation the output states read $\rho_{RA} = (\mathbb{I}_R \otimes \mathcal{E}_A) |\psi\rangle \langle \psi|_{RA}$ and $\sigma_{AB} = (\mathbb{I}_R \otimes \mathcal{F}_A) |\psi\rangle \langle \psi|_{RA}$. When $\mathcal{E}_A = \mathcal{F}_A$ we have $\sigma_{AB} = \rho_{AB}$, whose purifications, by the freedom of purification, must be equal up to a unitary operation U' on system E , as shown in Fig. A.1.

Measurement

Following Everett's interpretation of quantum mechanics [74], a measurement can be seen as a physical process similar to other quantum operations [39]. In a measurement the probe system A is entangled with an ancillary system P representing the pointer of the measurement apparatus.

Interestingly the unitary representation (A.9) of a quantum operation gives the physical model of the measurement that we need. The only difference is that now instead of discarding the auxiliary system as in Fig. A.1 (considered as lost into the environment), the state of the auxiliary system is accessible and gives the result of the measurement. Tracing the outgoing system A we obtain the probability distribution of the measurement result,

$$p(m) = \text{Tr}_A [E_m \rho_A E_m^\dagger]. \quad (\text{A.11})$$

If we apply a quantum non-demolition measurement, the state of system A conditioned on the pointer giving the result m reads,

$$\rho'_A = E_m \rho_A E_m^\dagger / p(m). \quad (\text{A.12})$$

Finally, the *completeness equation* of the measurement operators results from the completeness relation of the quantum operation. We have then shown that this physical model of measurement satisfies the three conditions of the measurement postulate of quantum mechanics.

Macroscopic measurement apparatus In a real measurement, in order to be readable by an human, the pointer system is composed by a macroscopic amount of particles which are entangled with the probe system as shown in Fig. A.2. The experimentalist having only access to the macroscopic apparatus, the outgoing quantum system is traced out, generating the randomness of the quantum measurement.

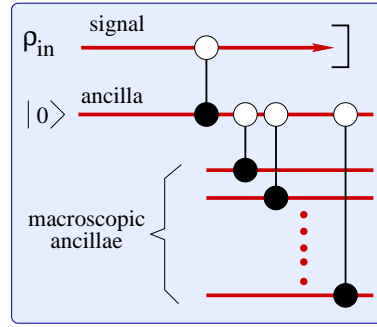


Figure A.2: The homodyne measurement can be seen as a C-NOT interaction between the quantum signal and a first ancilla ($|0\rangle$) which is followed by an amplification process, or "classicization", by entangling A with the rest of the apparatus ancillae through C-NOT gates. This entanglement is responsible for the randomness in the outcome.

Appendix B

Partial Measurement of a Bipartite Gaussian State

Consider a gaussian bipartite state ρ_{AB} defined by its covariance matrix

$$\gamma_{AB} = \begin{bmatrix} \gamma_A & \sigma_{AB} \\ \sigma_{AB}^T & \gamma_B \end{bmatrix}. \quad (\text{B.1})$$

and its mean (d_A, d_B) . If we apply a partial measurement measurement by projecting system B into a pure Gaussian state ψ of covariance matrix γ_M and mean m the Wigner function of the final system A reads,

$$W_\rho(r_A)' \propto \int dr_B W_\rho(r_A, r_B) W_\psi(r_B), \quad (\text{B.2})$$

where the proportionality comes from the fact that we do not pay attention the the probabilities as we are only interested on the conditional state of system A . It is convenient to deal with characteristic functions which are Fourier transform of the Wigner functions, $\chi(x) = \int W(r) \exp(ixr) dr$,

$$\chi_{\rho'}(x_A) \propto \int dx_B \chi_\rho(x_A, x_B) \chi_\psi(-x_B) \quad (\text{B.3})$$

Using the definition of the characteristic function of a Gaussian state (2.24) we obtain,

$$\begin{aligned} \chi'_\rho(x_A) &\propto \int \exp \left[-1/4(x_A^T \gamma_A x_A + x_A^T \sigma_{AB} x_B + x_B^T \sigma_{AB}^T x_A + x_B^T (\gamma_B + \gamma_M) x_B) \right] \\ &\times \exp[i d_A^T x_A + i(d_B - m)^T x_B] dx_B \end{aligned} \quad (\text{B.4})$$

which after rearranging and using the change of variable $z_B = x_B + (\gamma_B + \gamma_M)^{-1} \sigma_{AB}^T x_A$,

$$\begin{aligned} \chi'_\rho(x_A) &\propto \exp \left[-1/4(x_A^T (\gamma_A - \sigma_{AB} (\gamma_B + \gamma_M)^{-1} \sigma_{AB}^T) x_A) \right] \\ &\times \exp[i d_A^T x_A - i(d_B - m)^T (\gamma_B + \gamma_M)^{-1} \sigma_{AB}^T x_A] \\ &\times \int dz_B \exp \left[-1/4(z_B^T (\gamma_B + \gamma_M) z_B) \right] \exp[i(d_B - m)^T z_B] \end{aligned}$$

where the integral contributes only with a constant real number. Finally we see that the final state of system A is a Gaussian state of covariance matrix,

$$\gamma'_A = \gamma_A - \sigma_{AB} (\gamma_B + \gamma_M)^{-1} \sigma_{AB}^T \quad (\text{B.5})$$

and mean

$$d'_A = d_A + \sigma_{AB} (\gamma_B + \gamma_M)^{-1} (m - d_B). \quad (\text{B.6})$$

Interestingly the covariance matrix does not depend on the values of the measurement result m .

Appendix C

Properties of $t^{\hat{n}}$

Propagation of $t^{\hat{n}}$

In order to propagate the operator $t^{\hat{n}}$ to the right in 3.2 we use the relations

$$t^{\hat{n}} e^{\alpha^* \hat{a}} = e^{\alpha^* \hat{a}/t} t^{\hat{n}}, \quad t^{\hat{n}} e^{\alpha \hat{a}^\dagger} = e^{t\alpha \hat{a}^\dagger} t^{\hat{n}}, \quad (\text{C.1})$$

that we proof below. Using the Taylor development of $t^{\hat{n}}$

$$t^{\hat{n}} = e^{\hat{n} \log t} = \log t \left[1 + \hat{n}_t + \frac{\hat{n}_t^2}{2!} + \dots \right], \quad (\text{C.2})$$

with $\hat{n}_t = \hat{n} \log t$ and the commutator of $[\hat{n}, \hat{a}^\dagger] = \hat{a}^\dagger$, we obtain

$$\hat{n} \hat{a}^\dagger = \hat{a}^\dagger (\hat{n} + 1) \quad \text{and} \quad \hat{n} \hat{a} = \hat{a} (\hat{n} - 1), \quad (\text{C.3})$$

it is easy to obtain

$$t^{\hat{n}} \hat{a}^\dagger = \hat{a}^\dagger t^{\hat{n}+1} \quad \text{and} \quad t^{\hat{n}} \hat{a} = \hat{a} t^{\hat{n}-1}. \quad (\text{C.4})$$

Using the Taylor development of $e^{\alpha \hat{a}^\dagger}$

$$e^{\alpha \hat{a}^\dagger} = 1 + \alpha \hat{a}^\dagger + \frac{(\alpha \hat{a}^\dagger)^2}{2!} + \dots \quad (\text{C.5})$$

and equation (C.4) it is easy to show

$$t^{\hat{n}} e^{\alpha \hat{a}^\dagger} = e^{t\alpha \hat{a}^\dagger} t^{\hat{n}}. \quad (\text{C.6})$$

The proof for \hat{a} is similar to the previous one, where we use the commutator $[\hat{n}, \hat{a}] = -\hat{a}$ which gives finally gives the first equation of C.1.

Effect on $S(s_{in})|0\rangle$

$$t^{N\hat{n}} S(s)|0\rangle \sim \sum_{k=0}^{\infty} \frac{\sqrt{(2k)!}}{2^k k!} [\tanh(s_{in})]^k t^{N\hat{n}} |2k\rangle, \quad (\text{C.7})$$

$$\sim t^{2N} \sum_{k=0}^{\infty} \frac{\sqrt{(2k)!}}{2^k k!} [\tanh(s_{in})]^k |2k\rangle, \quad (\text{C.8})$$

which is exactly a squeezed state with a new squeezing factor $\tanh(s) = t^{2N} \tanh(s_{in})$.

Appendix D

Wigner Representation of Photon Subtraction

After applying a projector $\mathbb{I}_A \otimes X_B$ over a bipartite quantum state ρ_{AB} , the density matrix $\rho_{A,\text{out}}$ of mode A after a successful projection reads,

$$\rho_{A,\text{out}} = \text{Tr}_B[\rho_{AB}(\mathbb{I}_A \otimes X_B)]/P, \quad (\text{D.1})$$

where $P = \text{Tr}_{AB}[\rho_{AB}(\mathbb{I}_A \otimes X_B)]$ is the probability of a successful projection of the target state. The trace of the product of two operators can be evaluated by integrating the product of their Wigner representations over the phase space. Then the output Wigner function reads,

$$W_{\text{out}}(r)P = 2\pi \int W_\rho(r_A, r_B)W_X(r_B)dr_B, \quad (\text{D.2})$$

where P is the probability of a successful projection of the target state. In the following we will use the notation $W(r; \Gamma, d)$ for a Gaussian Wigner function with first mean d and covariance matrix $\gamma = \Gamma^{-1}$.

Partial trace: $X = \mathbb{I}$

When we project on the identity $X = \mathbb{I}$ we obtain,

$$W_{\text{out}}(r_A)P = 2\pi \int W(r_A, r_B; \Gamma_{AB}, d_{AB})dr_B = \frac{\sqrt{\det \Gamma_{AB}}}{\pi} \int e^{-Y} dr_B \quad (\text{D.3})$$

where Y reads,

$$Y = z_A^T \Gamma_A z_A + z_A^T \sigma z_B + z_B^T \sigma^T z_A + z_B^T \Gamma_B z_B, \quad (\text{D.4})$$

$z_{A(B)} = r_{A(B)} - d_{A(B)}$, and

$$\Gamma_{AB} = \begin{bmatrix} \Gamma_A & \sigma \\ \sigma^T & \Gamma_B \end{bmatrix}. \quad (\text{D.5})$$

Notice that Y can be rewritten as

$$Y = z_A^T \Gamma_A z_A + \underbrace{(z_B + \Gamma_B^{-1} \sigma^T z_A)^T \Gamma_B (z_B + \Gamma_B^{-1} \sigma^T z_A)}_{=Z} - z_A^T \sigma \Gamma_B^{-1} \sigma^T z_A \quad (\text{D.6})$$

we obtain,

$$\begin{aligned} W_{out}(r_A) &= \frac{\sqrt{\det \Gamma_{AB}}}{\pi} e^{-z_A^T (\Gamma_A - \sigma \Gamma_B^{-1} \sigma^T) z_A} \underbrace{\int e^{-Z} dr_B}_{1/\sqrt{\Gamma_B}} \\ &= \sqrt{\frac{\det \Gamma_{AB}}{\det \Gamma_B \det \Gamma_1}} W(r_A; \Gamma_1, d_1) \end{aligned} \quad (D.7)$$

where $\Gamma_1 = \Gamma_A - \sigma \Gamma_B^{-1} \sigma^T$ and $d_1 = d_A$. The Wigner function being normalized and the probability of projecting on the identity being trivially $P = 1$ we must have

$$\sqrt{\frac{\det \Gamma_{AB}}{\det \Gamma_B \det \Gamma_1}} = 1, \quad (D.8)$$

which holds as Γ_1 is the Schur complement of Γ_B of Γ_{AB} .

Projection on vacuum: $X = |0\rangle\langle 0|$

When we project on vacuum $X = |0\rangle\langle 0|$ we obtain,

$$\begin{aligned} W_{out}(r_A)P &= 2\pi \int W(r_A, r_B; \Gamma_{AB}, d_{AB}) W(r_A, r_B; \mathbb{I}, 0) dr_B \\ &= \frac{\sqrt{\det \Gamma_{AB}}}{\pi} \int e^{-Y} dr_B \end{aligned} \quad (D.9)$$

where Y reads,

$$Y = z_A^T \Gamma_A z_A + z_A^T \sigma z_B + z_B^T \sigma^T z_A + \underbrace{z_B^T \Gamma_B z_B + r_B^T \mathbb{I} r_B}_{=Z}, \quad (D.10)$$

where $z_{A(B)} = r_{A(B)} - d_{A(B)}$. Notice that Z can be rewritten as,

$$\begin{aligned} Z &= (r_B - (\Gamma_B + \mathbb{I})^{-1} \Gamma_B d_B)^T (\Gamma_B + \mathbb{I}) \underbrace{(r_B - (\Gamma_B + \mathbb{I})^{-1} \Gamma_B d_B)}_{\tilde{z}_B} \\ &\quad - d_B^T (\Gamma_B (\Gamma_B + \mathbb{I})^{-1} \Gamma_B - \Gamma_B) d_B. \end{aligned} \quad (D.11)$$

Then using $z_B = \tilde{z}_B + [(\Gamma_B + \mathbb{I})^{-1} \Gamma_B - \mathbb{I}] d_B$ and $(\Gamma_B + \mathbb{I})^{-1} \Gamma_B - \mathbb{I} = -(\Gamma_B + \mathbb{I})^{-1} Y$ reads,

$$\begin{aligned} Y &= z_A^T \Gamma_A z_A + [z_A^T \sigma \tilde{z}_B - z_A^T \sigma (\Gamma_B + \mathbb{I})^{-1} d_B] \\ &\quad + [\tilde{z}_B^T \sigma^T z_A - d_B^T \sigma (\Gamma_B + \mathbb{I})^{-1} z_A] \\ &\quad + \tilde{z}_B^T (\Gamma_B + \mathbb{I}) \tilde{z}_B + d_B^T \Gamma_B (\Gamma_B + \mathbb{I})^{-1} d_B, \end{aligned} \quad (D.12)$$

which can be rewritten using $\chi = \Gamma_A - \sigma (\Gamma_B + \mathbb{I})^{-1} \sigma^T$ and $\xi = \chi^{-1} \sigma (\Gamma_B + \mathbb{I})^{-1} d_B$

$$\begin{aligned} Y &= (z_A - \xi)^T \chi (z_A - \xi) - \xi^T \chi \xi + d_B^T \Gamma_B (\Gamma_B + \mathbb{I})^{-1} d_B \\ &\quad + \underbrace{\left[\tilde{z}_B + (\Gamma_B + \mathbb{I})^{-1} \sigma^T z_A \right]^T (\Gamma_B + \mathbb{I}) \left[\tilde{z}_B + (\Gamma_B + \mathbb{I})^{-1} \sigma^T z_A \right]}_W \end{aligned} \quad (D.13)$$

Then the integral (D.9) using $\int e^{-W} dr_B = \sqrt{\det(\Gamma_B + \mathbb{I})}$ reads,

$$W_{out}(r_A)P = 2 \sqrt{\frac{\det \Gamma_{AB}}{\det(\Gamma_B + \mathbb{I}) \det \Gamma_\chi}} e^{-d_B^T M d_B} W(r_A; \chi, \xi) \quad (D.14)$$

where $M = \Gamma_B(\Gamma_B + \mathbb{I})^{-1} - (\Gamma_B + \mathbb{I})^{-1}\sigma^T\chi^{-1}\sigma(\Gamma_B + \mathbb{I})^{-1}$. The Wigner function being normalized the probability of projecting on the vacuum reads,

$$P = 2\sqrt{\frac{\det \Gamma_{AB}}{\det(\Gamma_B + \mathbb{I}) \det \Gamma_\chi}}. \quad (\text{D.15})$$

Photon Subtraction

One uses the fact that the POVM element characterizing the photon subtraction ($\Pi_{1,B}$) is the difference of two POVM, the identity \mathbb{I}_B and the projection on vacuum $|0\rangle\langle 0|_B$,

$$W_{\Pi_1}(r) = \mathbb{I} - |0\rangle\langle 0|_B = \frac{1}{2\pi} - \frac{1}{\pi}e^{-x^2-p^2}. \quad (\text{D.16})$$

The Wigner function of mode A can be written as a linear combination of two Gaussian functions, namely

$$W(r)P = C_1W_G(r; \Gamma_1, d_1) + C_2W_G(r; \Gamma_2, d_2), \quad (\text{D.17})$$

where P is the probability of successful generation of the target state. The correlation matrix Γ_1 and the displacement d_1 appearing in the first term on the right-hand side of Eq. (3.31) are given by

$$\begin{aligned} \Gamma_1 &= \Gamma_A - \sigma(\Gamma_B + \mathbb{I})^{-1}\sigma^T, \\ d_1 &= d_A, \\ C_1 &= 1. \end{aligned} \quad (\text{D.18})$$

Similarly, the formulas for the parameters of the second term read

$$\begin{aligned} \Gamma_2 &= \Gamma_A - \sigma(\Gamma_B + \mathbb{I})^{-1}\sigma^T, \\ d_2 &= d_A + \Gamma_2^{-1}\sigma(\Gamma_B + \mathbb{I})^{-1}d_B, \\ C_2 &= -2\sqrt{\frac{\det(\Gamma_{AB})}{\det(\Gamma_2)\det(\Gamma_B + \mathbb{I})}}\exp[-d_B^T M d_B], \end{aligned} \quad (\text{D.19})$$

where

$$M = \Gamma_B(\Gamma_B + \mathbb{I})^{-1} - (\Gamma_B + \mathbb{I})^{-1}\sigma^T\Gamma_2^{-1}\sigma(\Gamma_B + \mathbb{I})^{-1}. \quad (\text{D.20})$$

Since all the Wigner functions are normalized, the probability of a successful photon subtraction reads $P = C_1 + C_2$.

Appendix E

Wigner Function from the Fock Basis

A single mode quantum state of light can be written using the Fock state representation as,

$$\rho = \sum_{m,n} \rho_{m,n} |m\rangle\langle n|. \quad (\text{E.1})$$

The Wigner function using polar coordinates reads,

$$W(r, \theta) = \sum_{m,n} \rho_{m,n} W_{m,n}(r, \theta), \quad (\text{E.2})$$

where $(x = r \cos \theta, p = r \sin \theta)$ and $W_{m,n}(r, \theta)$ is the Wigner function of the operator $|m\rangle\langle n|$ which reads,

$$\begin{aligned} W_{m,n}(r, \theta) &= \frac{(-1)^n}{\pi} \left[\frac{n!}{m!} \right]^{1/2} e^{i(m-n)\theta} (\sqrt{2}r)^{m-n} e^{-r^2} L_n^{m-n}(2r^2) \quad m \geq n \\ W_{m,n} &= W_{n,m}^* \end{aligned} \quad (\text{E.3})$$

as shown in [86], where $L_n^{m-n}(x)$ is a Laguerre polynomial,

$$L_n^k(x) = \sum_{m=0}^n (-1)^m \frac{(n+k)!}{(n-m)!(k+m)!m!} x^m. \quad (\text{E.4})$$

Appendix F

Ideal Photon Subtraction

The initial bipartite two-mode squeezed vacuum state reads,

$$|\psi\rangle_{in} = \sqrt{1-\lambda^2} \sum_{n=0}^{\infty} \lambda^n |n\rangle_A \otimes |n\rangle_B, \quad (\text{F.1})$$

where the state is normalized $\langle\psi|\psi\rangle = 1$ as we have,

$$S_0(x) = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}. \quad (\text{F.2})$$

Two Photon Subtractions

The conditional photon subtraction on each mode can be described by the non-unitary operator (with the same transmittance T on both beamsplitters),

$$\hat{X}_A \otimes \hat{X}_B = t^{\hat{n}_A} r \hat{a}_A \otimes t^{\hat{n}_B} r \hat{a}_B, \quad (\text{F.3})$$

where $t = \sqrt{T}$, as shown in Chapter 4. The conditional generated state can be written, using the relations $\hat{a}|n\rangle = \sqrt{n}|n-1\rangle$, $t^{\hat{n}}\hat{a} = \hat{a}t^{\hat{n}-1}$ and $t^{\hat{n}}|n\rangle = t^n|n\rangle$ (see Appendix C),

$$\sqrt{p_2}|\psi\rangle_{out} = \sqrt{1-\lambda^2}(1-T)\lambda \sum_{n=0}^{\infty} (T\lambda)^n (n+1) |n\rangle_A \otimes |n\rangle_B, \quad (\text{F.4})$$

where p_2 is the probability of success of the double photon subtraction.

First we determine the state conditioned on a successful subtraction,

$$|\psi\rangle_{out} \propto \sum_{n=0}^{\infty} (n+1) (T\lambda)^n |n\rangle_A \otimes |n\rangle_B. \quad (\text{F.5})$$

In order to normalize it we need to calculate $\sum_{n=0}^{\infty} (T\lambda)^{2n} (n+1)^2$. Deriving among x the equation (F.2) we obtain,

$$S_1(x) = \frac{d}{dx} S_0(x) = \sum_{n=0}^{\infty} (n+1) x^n = \frac{1}{(1-x)^2}. \quad (\text{F.6})$$

Applying a second derivative we get,

$$S_2(x) = \frac{d^2}{dx^2} S_0(x) = \sum_{n=0}^{\infty} (n+2)(n+1) x^n = \frac{2}{(1-x)^3}. \quad (\text{F.7})$$

Combining equations (F.6) and (F.7) we obtain

$$\sum_{n=0}^{\infty} (T\lambda)^{2n} (n+1)^2 = S_2(T^2\lambda^2) - S_1(T^2\lambda^2) = \frac{1 + T^2\lambda^2}{(1 - T^2\lambda^2)^3}. \quad (\text{F.8})$$

Finally the final state reads,

$$|\psi\rangle_{out} = \sqrt{\frac{(1 - T^2\lambda^2)^3}{1 + T^2\lambda^2}} \sum_{n=0}^{\infty} \lambda^n (n+1) |n\rangle_A \otimes |n\rangle_B, \quad (\text{F.9})$$

and the probability of successful photon subtraction reads,

$$p_2 = (1 - T)^2 \lambda^2 (1 - \lambda^2) \frac{1 + T^2\lambda^2}{(1 - T^2\lambda^2)^3}. \quad (\text{F.10})$$

Four Photon Subtractions

The double conditional photon subtraction is model by applying the non-unitary operator $\hat{X}_A^2 \otimes \hat{X}_B^2$. The output state reads,

$$\sqrt{p_4} |\psi\rangle_{out} = \sqrt{1 - \lambda^2} (1 - T)^2 T \lambda^2 \sum_{n=0}^{\infty} (n+2)(n+1) (T^2\lambda)^n |n\rangle_A \otimes |n\rangle_B, \quad (\text{F.11})$$

where p_4 is the probability of success of the four photon subtraction. First we determine the state conditioned on a successful subtraction,

$$|\psi\rangle_{out} \propto \sum_{n=0}^{\infty} (n+2)(n+1) (T^2\lambda)^n |n\rangle_A \otimes |n\rangle_B. \quad (\text{F.12})$$

In order to normalize it we need to calculate $\sum_{n=0}^{\infty} (n+2)^2 (n+1)^2 (T^2\lambda)^{2n}$ we need to define

$$S_3(x) = \frac{d^3}{dx^3} S_0(x) = \sum_{n=0}^{\infty} \frac{(n+3)!}{n!} x^n = \frac{6}{(1-x)^4}, \quad (\text{F.13})$$

and

$$S_4(x) = \frac{d^4}{dx^4} S_0(x) = \sum_{n=0}^{\infty} \frac{(n+4)!}{n!} x^n = \frac{24}{(1-x)^5}. \quad (\text{F.14})$$

Combining equations (F.7), (F.13) and (F.14) we obtain

$$\sum_{n=0}^{\infty} (n+2)^2 (n+1)^2 (T^2\lambda)^{2n} = \quad (\text{F.15})$$

$$= S_4(T^4\lambda^2) - 4S_3(T^4\lambda^2) + 2S_2(T^4\lambda^2) = 4 \frac{1 + 4T^4\lambda^2 + T^8\lambda^4}{(1 - T^4\lambda^2)^5}. \quad (\text{F.16})$$

Finally the final state reads,

$$|\psi\rangle_{out} = \frac{1}{2} \sqrt{\frac{(1 - T^4\lambda^2)^5}{1 + 4T^4\lambda^2 + T^8\lambda^4}} \sum_{n=0}^{\infty} (n+2)(n+1) \lambda^n |n\rangle_A \otimes |n\rangle_B, \quad (\text{F.17})$$

and the probability of successfully subtracting four photons reads,

$$p = 4T^2(1 - T)^4 \lambda^4 (1 - \lambda^2) \frac{1 + 4T^4\lambda^2 + T^8\lambda^4}{(1 - T^4\lambda^2)^5}. \quad (\text{F.18})$$

Appendix G

Calculation of G

In order to simplify the calculation,

$$G = \int_0^\infty \int_0^\infty e^{-(ax^2+by^2+2cxy)} dx dy, \quad (\text{G.1})$$

we use polar coordinates ($x = r \cos \theta, y = r \sin \theta$) giving

$$G = \int_0^\infty \int_0^{\pi/2} e^{-r^2(a \cos^2 \theta + b \sin^2 \theta + 2c \sin \theta \cos \theta)} r dr d\theta. \quad (\text{G.2})$$

Applying the following change of variables $z = \alpha r^2$ we obtain,

$$\int_0^\infty e^{-\alpha r^2} r dr = \frac{1}{2\alpha} \int_0^\infty e^{-z} dz = \frac{1}{2\alpha}, \quad (\text{G.3})$$

which gives us,

$$G = \frac{1}{2} \int_0^{\pi/2} \frac{1}{a \cos^2 \theta + b \sin^2 \theta + 2c \sin \theta \cos \theta} d\theta. \quad (\text{G.4})$$

Then G can be rewritten as,

$$G = \frac{1}{2} \int_0^{\pi/2} \frac{1/\cos^2 \theta}{a + b \tan^2 \theta + 2c \tan \theta} d\theta, \quad (\text{G.5})$$

which applying the following change of variables $w = \tan \theta$ reads (see [2] for the integral),

$$G = \frac{1}{2} \int_0^\infty \frac{1}{a + bw^2 - 2cw} dw \quad (\text{G.6})$$

$$= \frac{1}{2\sqrt{ab-c^2}} \left[\arctan \frac{c+bw}{\sqrt{ab-c^2}} \right]_0^\infty \quad (\text{G.7})$$

$$= \frac{1}{2\sqrt{ab-c^2}} \left(\frac{\pi}{2} - \arctan \frac{c}{\sqrt{ab-c^2}} \right). \quad (\text{G.8})$$

Appendix H

Incremental Proportionality of Mutual Information

We want to proof the following statement.

By sending R bits of communication we can increase the correlations at most by R shared perfect random bits.

At the beginning Alice has two systems X and X' , see Fig. H.1, and Bob just one Y , their correlations are measured by $H(X, X':Y)$. After Alice has sent the system X' to Bob their correlations read $H(X: X', Y)$. The final correlations reads,

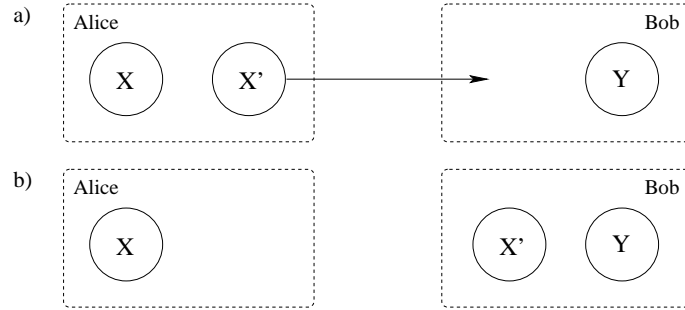


Figure H.1: a) At the beginning Alice has two system X and X' , and Bob just one Y . Later Alice sends the system X' to Bob. b) At the end Alice has system X and Bob has X' and Y .

$$H(X:X', Y) = H(X) + H(X'Y) - H(X, X', Y) \quad (\text{H.1})$$

which can be upperbounded using the subadditivity of the entropy,

$$H(X:X', Y) \leq H(X) + H(X') + H(Y) - H(X, X', Y). \quad (\text{H.2})$$

Then using $H(X) \leq H(X, X')$ we obtain,

$$H(X:X', Y) \leq [H(X, X') + H(Y) - H(X, X', Y)] + H(X') \quad (\text{H.3})$$

$$= H(X, X':Y) + H(X'). \quad (\text{H.4})$$

We see that sending a signal X' taking values from an alphabet of size d the maximum increase of correlations is $\log d$, which in the case of a binary alphabet is just 1 bit.

Appendix I

Distance between Purifications

Trace Distance and Fidelity

Two different measures are the most commonly used to measure the closeness of two quantum states (ρ and σ), the trace distance

$$D(\rho, \sigma) = \frac{1}{2} \|\rho - \sigma\|_1 \quad (\text{I.1})$$

and the fidelity

$$F(\rho, \sigma) = \text{Tr} \sqrt{\rho^{1/2} \sigma \rho^{1/2}}. \quad (\text{I.2})$$

Uhlmann's Theorem [189, 136]

Consider ρ and σ are states of a quantum system Q . Introduce a second register R which is a copy of Q . Then

$$F(\rho, \sigma) = \max_{|\psi\rangle, |\varphi\rangle} |\langle\psi|\varphi\rangle| \quad (\text{I.3})$$

where the maximization is over all purifications $|\psi\rangle$ of ρ and $|\varphi\rangle$ of σ into RQ .

Closeness of the Purifications

Consider two bipartite quantum states shared by Alice and Bob, ρ_{AB} and σ_{AB} . If the two quantum state are close,

$$\|\rho_{AB} - \sigma_{AB}\|_1 \leq \epsilon. \quad (\text{I.4})$$

Using the lower bound of the relation [136]

$$1 - F(\rho, \sigma) \leq D(\rho, \sigma) \leq \sqrt{1 - F(\rho, \sigma)^2} \quad (\text{I.5})$$

we obtain

$$F(\rho_{AB}, \sigma_{AB}) \geq 1 - \epsilon/2. \quad (\text{I.6})$$

Using Uhlmann's theorem we can show that there exists one purification Ψ_{ABE} of ρ_{AB} and Φ_{ABE} of σ_{AB} which satisfies

$$F(\Psi_{ABE}, \Phi_{ABE}) \geq 1 - \epsilon/2 \quad (\text{I.7})$$

Then using the upperbound (I.5) we obtain,

$$\|\Psi_{ABE} - \Phi_{ABE}\|_1 \leq 2\sqrt{1 - (1 - \epsilon/2)^2} \leq 2\sqrt{\epsilon}. \quad (\text{I.8})$$

□

Appendix J

Detail of Calculation of Section (8.4)

Calculation of equation (8.92)

We develop

$$\rho = \frac{r'^2}{s'^2 + t'^2} \quad (\text{J.1})$$

using the definition of r' , s' et t' in

$$S_P = \frac{1}{d\sqrt{T}} \begin{pmatrix} bt - cs & cr - at & as - br \\ \underbrace{\frac{\sqrt{\chi}(s \sin \theta - t \cos \theta)}{\sqrt{\chi}(c \cos \theta - b \sin \theta)}}_{r'} & \underbrace{\frac{t - r\sqrt{\chi} \sin \theta}{a\sqrt{\chi} \sin \theta - c}}_{s'} & \underbrace{\frac{r\sqrt{\chi} \cos \theta - s}{b - a\sqrt{\chi} \cos \theta}}_{t'} \end{pmatrix} \quad (\text{J.2})$$

that gives us,

$$\chi(c \cos \theta - b \sin \theta)^2 = \rho [(a\sqrt{\chi} \sin \theta - c)^2 + (b - a\sqrt{\chi} \cos \theta)^2]. \quad (\text{J.3})$$

This equation can be simplified using the definition of a , b and c ,

$$a = u\sqrt{\rho} \quad (\text{J.4})$$

$$b = u \sin \xi \quad (\text{J.5})$$

$$c = u \cos \xi \quad (\text{J.6})$$

giving,

$$\chi \cos^2 (\xi + \theta) = \rho(\rho\chi + 1 - 2\sqrt{\rho\chi} \sin (\xi + \theta)). \quad (\text{J.7})$$

Using the identity $\cos^2 (\xi + \theta) + \sin^2 (\xi + \theta) = 1$, we obtain

$$(\chi - \rho) \cos^2 (\xi + \theta) = \rho(\rho\chi + \sin^2 (\xi + \theta) - 2\sqrt{\rho\chi} \sin (\xi + \theta)) \quad (\text{J.8})$$

which simplifies to

$$\left(\frac{\chi - \rho}{\rho} \right) \cos^2 (\xi + \theta) = (\sin (\xi + \theta) - \sqrt{\rho\chi})^2 \quad (\text{J.9})$$

Calculation of equation (8.93)

Developping

$$(cr - at)^2 + (as - br)^2 = r^2(b^2 + c^2) + a^2(t^2 + s^2) - 2ar(ct + bs) = d^2 T^2 \chi \quad (\text{J.10})$$

using the definition of a, b and c we obtain

$$r^2 + \rho(t^2 + s^2) - 2r\sqrt{\rho}(t \cos \xi + s \sin \xi) = \frac{d^2 T^2 \chi}{u^2}. \quad (\text{J.11})$$

Using

$$bt - cs = dT \quad (\text{J.12})$$

and the definition of b and c we obtain

$$ut \sin \xi - us \cos \xi = dT \quad (\text{J.13})$$

that squaring and adding $(t \cos \xi + s \sin \xi)^2$ on both sides gives,

$$t^2 + s^2 = \left(\frac{dT}{u} \right)^2 + (t \cos \xi + s \sin \xi)^2 \quad (\text{J.14})$$

Replacing (J.14) into (J.11) gives

$$\rho(t \cos \xi + s \sin \xi)^2 - 2r\sqrt{\rho}(t \cos \xi + s \sin \xi) + r^2 = \frac{d^2 T^2}{u^2}(\chi - \rho) \quad (\text{J.15})$$

that can be simplified to

$$\left[r - \sqrt{\rho}(t \cos \xi + s \sin \xi) \right]^2 = \left(\frac{dT}{u} \right)^2 (\chi - \rho) \quad (\text{J.16})$$

writing s as a function of t using (J.13) we obtain

$$\left[r - t \frac{\sqrt{\rho}}{\cos \xi} + \frac{dT\sqrt{\rho}}{u} \tan \xi \right]^2 = \left(\frac{dT}{u} \right)^2 (\chi - \rho) \quad (\text{J.17})$$

We can obtain a second equation on r and t by using the definition of the determinant $d = \det S_X$,

$$d = (bt - cs) + \sqrt{\chi} \cos \theta (rc - ta) + \sqrt{\chi} \sin \theta (as - rb) \quad (\text{J.18})$$

Replacing $(bt - cs)$ by dT and using the definition of a, b and c we obtain,

$$r \cos(\theta + \xi) + \sqrt{\rho}(s \sin \theta - t \cos \theta) = \frac{d}{u} \frac{(1 - T)}{\sqrt{\chi}} \quad (\text{J.19})$$

Writing s as a function of t using (J.13) we obtain

$$r \cos(\theta + \xi) + t\sqrt{\rho}(\sin \theta \tan \xi - \cos \theta) = \frac{d}{u} \left[\frac{(1 - T)}{\sqrt{\chi}} + T\sqrt{\rho} \frac{\sin \theta}{\cos \xi} \right] \quad (\text{J.20})$$

that can be simplified to

$$r - t \frac{\sqrt{\rho}}{\cos \xi} = \frac{d}{u \cos(\theta + \xi)} \left[\frac{(1 - T)}{\sqrt{\chi}} + T\sqrt{\rho} \frac{\sin \theta}{\cos \xi} \right]. \quad (\text{J.21})$$

Injecting equation (J.21) into (J.17) finally gives,

$$\left(\frac{\chi - \rho}{\rho} \right) \cos^2(\xi + \theta) = \left(\sin(\xi + \theta) + \frac{1 - T}{T\sqrt{\rho\chi}} \right)^2. \quad (\text{J.22})$$

Bibliography

- [1] A. Abeyesinghe, I. Devetak, P. Hayden, and A. Winter. The mother of all protocols: Restructuring quantum information's family tree. *arXiv:quant-ph/0606225*, 2006.
- [2] M. Abramowitz and I. A. Stegun. *Handbook of Mathematical Functions*. Dover, New York, 1991.
- [3] U. L. Andersen, O. Glöckl, S. Lorenz, G. Leuchs, and R. Filip. Experimental demonstration of continuous variable quantum erasing. *Phys. Rev. Lett.* **93**, 100403, 2004.
- [4] U. L. Andersen, V. Josse, and G. Leuchs. Unconditional quantum cloning of coherent states with linear optics. *Phys. Rev. Lett.* **94**, 240503, 2005.
- [5] U. L. Andersen, M. Sabuncu, R. Filip, and G. Leuchs. Experimental demonstration of coherent state estimation with minimal disturbance. *Phys. Rev. Lett.* **96**, 020409, 2006.
- [6] M. Arndt, O. Nairz, J. Voss-Andreae, C. Keller, G. van der Zouw, and A. Zeilinger. Wave-particle duality of c60. *Nature* **401**, 680, 1999.
- [7] A. Aspect. Proposed experiment to test the nonseparability of quantum mechanics. *Phys. Rev. D* **14**, 1944, 1976.
- [8] A. Aspect, J. Dalibard, and G. Roger. Experimental test of Bell's inequalities using time-varying analyzers. *Phys. Rev. Lett.* **49**, 1804, 1982.
- [9] A. Aspect, P. Grangier, and G. Roger. Experimental tests of realistic local theories via Bell's theorem. *Phys. Rev. Lett.* **47**, 460, 1981.
- [10] A. Aspect, P. Grangier, and G. Roger. Experimental realization of Einstein-Podolsky-Rosen-Bohm gedankenexperiment: A new violation of Bell's inequalities. *Phys. Rev. Lett.* **49**, 91, 1982.
- [11] S. A. Babichev, B. Brezger, and A. I. Lvovsky. Remote preparation of a single-mode photonic qubit by measuring field quadrature noise. *Phys. Rev. Lett.* **92**, 047903, 2004.
- [12] S. A. Babichev, J. Ries, and A. I. Lvovsky. Quantum scissors: teleportation of single-mode optical states by means of a nonlocal single photon. *Europhys. Lett.* **64**, 1, 2003.
- [13] K. Banaszek and K. Wódkiewicz. Nonlocality of the Einstein-Podolsky-Rosen state in the Wigner representation. *Phys. Rev. A* **58**, 4345, 1998.
- [14] H. Bechmann-Pasquinucci and N. Gisin. Incoherent and coherent eavesdropping in the six-state protocol of quantum cryptography. *Phys. Rev. A* **59**, 4238, 1999.

- [15] J. S. Bell. On Einstein-Podolsky-Rosen paradox. *Physics (Long Island City, N.Y.)* **1**, 195, 1964.
- [16] J.S. Bell. *Speakable and Unspeakable in Quantum Mechanics*. CUP, Cambridge, 1988.
- [17] K. Bencheikh, J.A. Levenson, Ph. Grangier, and O. Lopez. Quantum nondemolition demonstration via repeated backaction evading measurements. *Phys. Rev. Lett.* **75**, 3422, 1995.
- [18] K. Bencheikh, C. Simonneau, and J. A. Levenson. Cascaded amplifying quantum optical taps: A robust noiseless optical bus. *Phys. Rev. Lett.* **78**, 34, 1995.
- [19] C. Bennett. Quantum cryptography using any two nonorthogonal states. *Phys. Rev. Lett.* **68**, 3121, 1992.
- [20] C. Bennett and G. Brassard. Proceeding of the ieee international conference on computers, systems and signal processing, bangalore, india. *IEEE, New York, p.* 175, 1984.
- [21] C. Bennett, C. Crépeau G. Brassard, R. Jozsa, A. Peres, and W. Wootters. Teleporting an unknown quantum state via dual classical and epr channels. *Phys. Rev. Lett.* **70**, 1899, 1993.
- [22] C. Bennett and S. J. Wiesner. Communication via one- and two-particle operators on einstein-podolsky-rosen states. *Phys. Rev. Lett.* **69**, 2881, 1992.
- [23] G. Björk, L.L. Sánchez-Soto, and J. Söderholm. Entangled-state lithography: Tailoring any pattern with a single state. *Phys. Rev. Lett.* **86**, 4516, 2001.
- [24] B. B. Blinov, D. L. Moehring, L. M. Duan, and C. Monroe. Observation of entanglement between a single trapped atom and a single photon. *Nature (London)* **428**, 153, 2004.
- [25] N. Bohr. On the constitution of atoms and molecules. *Philosophical Magazine* **26**, 1, 1913.
- [26] N. Bohr. The spectra of helium and hydrogen. *Nature* **92**, 231, 1914.
- [27] N. Bohr. Atomic structure. *Nature*, **107**, 104, 1921.
- [28] M. Born. Zur quantenmechanik der stoßvorgänge. *Z. Phys.* **37**, 863, 1926.
- [29] M. Born, W. Heisenberg, and P. Jordan. Zur quantenmechanik ii. *Z. Phys.* **35**, 557, 1925.
- [30] A.N. Boto, P. Kok, D.S. Abrams, S.L. Braunstein, C.P. Williams, and J.P. Dowling. Quantum interferometric optical lithography: Exploiting entanglement to beat the diffraction limit. *Phys. Rev. Lett.* **85**, 2733, 2000.
- [31] D. Bouwmeester, A. Ekert, and A. Zeilinger. *The Physics of Quantum Information*. Springer, Berlin, 2000.
- [32] W. P. Bowen, R. Schnabel, P. K. Lam, and T. C. Ralph. Experimental characterization of continuous-variable entanglement. *Phys. Rev. A* **69**, 012304, 2004.
- [33] S. L. Braunstein. Squeezing as an irreducible resource. *quant-ph/9904002*, 1999.

- [34] S. L. Braunstein, N. J. Cerf, S. Iblidir, P. van Loock, and S. Massar. Optimal cloning of coherent states with linear amplifier and beam splitter. *Phys. Rev. Lett.* **86**, 4938, 2001.
- [35] S. L. Braunstein and H. J. Kimble. Teleportation of continuous quantum variables. *Phys. Rev. Lett.* **80**, 869, 1998.
- [36] H.J. Briegel, W. Dür, J.I. Cirac, and P. Zoller. Quantum repeaters: The role of imperfect local operations in quantum communication. *Phys. Rev. Lett.* **81**, 5932, 1998.
- [37] D.E. Browne, J. Eisert, S. Scheel, and M.B. Plenio. Driving non-gaussian to gaussian states with linear optics. *Phys. Rev. A* **67**, 062320, 2003.
- [38] D. Bruss. Optimal eavesdropping in quantum cryptography with six states. *Phys. Rev. Lett.* **81**, 3018, 1998.
- [39] N. J. Cerf and C. Adami. Quantum mechanics of measurement. *quant-ph/9605002*, 1996.
- [40] N. J. Cerf and J. Fiurasek. Optical quantum cloning. *Progress in Optics*, 49, edited by E. Wolf, (Elsevier, Amsterdam, 2006), pp. 455-545., 2006.
- [41] N. J. Cerf and S. Iblidir. Optimal n-to-m cloning of conjugate quantum variables. *Phys. Rev. A* **62**, 040301(R), 2000.
- [42] N. J. Cerf and S. Iblidir. Phase conjugation of continuous quantum variable. *Phys. Rev. A* **64**, 032307, 2001.
- [43] N. J. Cerf, A. Ipe, and X. Rottenberg. Cloning of continuous quantum variable. *Phys. Rev. Lett.* **85**, 1754, 2000.
- [44] N. J. Cerf, M. Levy, and G. Van Assche. Quantum distribution of Gaussian keys using squeezed states. *Phys. Rev. A* **63**, 052311, 2001.
- [45] N.J. Cerf, J. Fiurásek, S. Iblidir, and S. Massar. Generation of large photon-number cat states using linear optics and quantum memory. *Proceedings of the 6th International Conference on Quantum Communication, Measurement and Computing*, eds. J. H. Shapiro and O. Hirota, p. 249, 2003.
- [46] Z.-B. Chen, J.-W. Pan, G. Hou, and Y.-D. Zhang. Maximal violation of Bell's inequalities for continuous variable systems. *Phys. Rev. Lett.* **88**, 040406, 2002.
- [47] J. Clausen, N. Hansen, L. Knoll, J. Mlynek, and D.G. Welsch. Conditional quantum-state engineering in repeated 2-photon down-conversion. *Applied Physics B-Lasers and Optics* **72**, 43, 2001.
- [48] J. F. Clauser, M.A. Horne, A. Shimony, and R.A. Holt. Proposed experiment to test local hidden-variable theories. *Phys. Rev. Lett.* **23**, 880, 1969.
- [49] P.T. Cochrane, T.C. Ralph, and G.J. Milburn. Teleportation improvement by conditional measurements on the two-mode squeezed vacuum. *Phys. Rev. A* **65**, 062306, 2002.
- [50] A. H. Compton. A quantum theory of the scattering of x-rays by light elements. *Phys. Rev.* **21**, 483, 1923.
- [51] T. M. Cover and J. A. Thomas. *Elements of Information Theory*. Wiley and Sons, 1991.

- [52] I. Csiszar and J. Korner. Broadcast channels with confidential messages. *IEEE Transaction on Information Theory*, **24**(3), 1978.
- [53] M. Dakna, T. Anhut, T. Opatrny, L. Knöll, and D.-G. Welsch. Generating Schrödinger-cat-like states by means of conditional measurements on a beam splitter. *Phys. Rev. A* **55**, 3184, 1997.
- [54] M. Dakna, J. Clausen, L. Knöll, and D.-G. Welsch. Erratum: Generation of arbitrary quantum states of traveling fields [phys. rev. a 59, 1658 (1999)]. *Phys. Rev. A* **60**, 726, 1999.
- [55] M. Dakna, J. Clausen, L. Knöll, and D.-G. Welsch. Generation of arbitrary quantum states of traveling fields. *Phys. Rev. A* **59**, 1658, 1999.
- [56] C. J. Davisson and L. H. Germer. The scattering of electrons by a single crystal of nickel. *Nature* **119**, 558, 1927.
- [57] L. de Broglie. Recherches sur la théorie des quanta. *Thesis (Paris)*, 1924, 1924.
- [58] I. Devetak, A. W. Harrow, and A. Winter. A family of quantum protocols. *Phys. Rev. Lett.* **93**, 230504, 2004.
- [59] I. Devetak and A. Winter. Relating quantum privacy and quantum coherence: An operational approach. *Phys. Rev. Lett.* **93**, 080501, 2004.
- [60] P. A. M. Dirac. *The principles of quantum mechanics*. Oxford university press, Oxford, 1930.
- [61] DiVincenzo, P. Shor, and J. Smolin. Quantum-channel capacity of very noisy channels. *Phys. Rev. A* **57**, 830, 1998.
- [62] L. M. Duan, J. I. Cirac, P. Zoller, and E. S. Polzik. Quantum communication between atomic ensembles using coherent light. *Phys. Rev. Lett.* **85**, 5643, 2000.
- [63] L. M. Duan, G. Giedke¹, J. I. Cirac, and P. Zoller. Inseparability criterion for continuous variable systems. *Phys. Rev. Lett.* **84**, 2722, 2000.
- [64] L. M. Duan and H. J. Kimble. Efficient engineering of multiatom entanglement through single-photon detections. *Phys. Rev. Lett.* **90**, 253601, 2003.
- [65] L.M. Duan, M.D. Lukin, J.I. Cirac, and P. Zoller. Long-distance quantum communication with atomic ensembles and linear optics. *Nature (London)* **414**, 413, 2001.
- [66] A. Einstein. Zur elektrodynamik bewegter körper. *Ann. D. Phys.* **17**, 891, 1905.
- [67] A. Einstein. Über einen die erzeugung und verwandlung des liches betreffenden heuristischen gesichtspunkt. *Ann. D. Phys.* **17**, 132, 1905.
- [68] A. Einstein. Die plancksche theorie der strahlung und die theorie der spezifischen wärme. *Ann. D. Phys.* **22**, 180, 1907.
- [69] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.* **47**, 777, 1935.
- [70] H.S. Eisenberg, J.F. Hodelin, G. Khoury, and D. Bouwmeester. Multiphoton path entanglement by nonlocal bunching. *Phys. Rev. Lett.* **94**, 090502, 2005.

- [71] J. Eisert, D. Browne, S. Scheel, and M. B. Plenio. Distillation of continuous-variable entanglement with optical means. *Annals of Physics (NY)* **311**, 431, 2004.
- [72] J. Eisert and M. Plenio. Introduction to the basics of entanglement theory in continuous-variable system. *arXiv:quant-ph/0312071*, 2003.
- [73] J. Eisert, S. Scheel, and M.B. Plenio. Distilling gaussian states with gaussian operations is impossible. *Phys. Rev. Lett.* **89**, 137903, 2002.
- [74] H. Everett. 'relative state' formulation of quantum mechanics. *Rev. Mod. Phys.* **29**, 454, 1957.
- [75] X.-L. Feng, X.-D. Li Z.-M. Zhang, S.-Q. Gong, and Z.-Z. Xu. Entangling distant atoms by interference of polarized photons. *Phys. Rev. Lett.* **90**, 217902, 2003.
- [76] R. Filip and Jr. L. Mišta. Violation of Bell's inequalities for a two-mode squeezed vacuum state in lossy transmission lines. *Phys. Rev. A* **66**, 044309, 2002.
- [77] R. Filip, P. Marek, and U.L. Andersen. Measurement-induced continuous-variable quantum interactions. *Rev. A* **71**, 042308, 1995.
- [78] J. Fiurasek. Optical implementation of continuous-variable quantum cloning machines. *Phys. Rev. Lett.* **86**, 4942, 2001.
- [79] J. Fiurášek. Conditional generation of n-photon entangled states of light. *Phys. Rev. A* **65**, 053818, 2002.
- [80] S. J. Freedman and J. F. Clauser. Experimental test of local hidden-variable theories. *Phys. Rev. Lett.* **28**, 938, 1972.
- [81] M. Freyberger, P. K. Aravind, M. A. Horne, and A. Shimony. Proposed test of Bell's inequality without a detection loophole by using entangled rydberg atoms. *Phys. Rev. A* **53**, 1232, 1996.
- [82] E. S. Fry, T. Walther, and S. Li. Proposal for a loophole-free test of the Bell inequalities. *Phys. Rev. A* **52**, 4381, 1995.
- [83] A. Furusawa, J.L. Sørensen, S.L. Braunstein, C.A. Fuchs, H.J. Kimble, and E.S. Polzik. Unconditional quantum teleportation. *Science* **282**, 706-709, 1998.
- [84] R. García-Patrón and N.J. Cerf. Unconditional optimality of Gaussian attacks against continuous-variable quantum key distribution. *Phys. Rev. Lett.* **97**, 190503, 2006.
- [85] R. García-Patrón, J. Fiurášek, N.J. Cerf, J. Wenger, R. Tualle-Brouiri, and Ph. Grangier. Proposal for a loophole-free Bell test using homodyne detection. *Phys. Rev. Lett.* **93**, 130409, 2004.
- [86] B. M. Garraway and P. L. Knight. Quantum phase distributions and quasidistributions. *Phys. Rev. A* **46**, 5346(R), 1992.
- [87] H. Geiger and E. Mardsen. On a diffuse reflection of the α -particles. *Proceedings of the Royal Society, Series A* **82**: 495-500, 1909.
- [88] C.C. Gerry, A. Benmoussa, and R. A. Campos. Nonlinear interferometer as a resource for maximally entangled photonic states: Application to interferometry. *Phys. Rev. A* **66**, 013804, 2002.

- [89] G. Giedke and J.I. Cirac. Characterization of Gaussian operations and distillation of Gaussian states. *Phys. Rev. A* **66**, 032316, 2002.
- [90] A. Gilchrist, P. Deuar, and M. D. Reid. Contradiction of quantum mechanics with local hidden variables for quadrature phase amplitude measurements. *Phys. Rev. Lett.* **80**, 3169, 1998.
- [91] A. Gilchrist, P. Deuar, and M. D. Reid. Contradiction of quantum mechanics with local hidden variables for quadrature phase measurements on pair-coherent states and squeezed macroscopic superpositions of coherent states. *Phys. Rev. A* **60**, 4259, 1999.
- [92] N. Gisin, G. Ribordy, W. Tittel, and H. Zbinden. Quantum cryptography. *Rev. Mod. Phys.* **74**, 145, 2002.
- [93] B. Groisman, S. Popescu, and A. Winter. Quantum, classical, and total amount of correlations in a quantum state. *Phys. Rev. A* **72**, 032317, 2005.
- [94] F. Grosshans, G. Van Assche, J. Wenger, R. Tualle-Brouri, N. J. Cerf, and P. Grangier. Quantum key distribution using gaussian-modulated coherent states. *Nature* **421**, 238, 2003.
- [95] F. Grosshans, N. J. Cerf, J. Wenger, R. Tualle-Brouri, and P. Grangier. Virtual entanglement and reconciliation protocols for quantum cryptography with continuous variables. *Quantum Inf. Comput.* **3**, 535, 2003.
- [96] F. Grosshans and P. Grangier. Continuous variable quantum cryptography using coherent states. *Phys. Rev. Lett.* **88**, 057902, 2002.
- [97] P. Hausladen, R. Jozsa, B. Schumacher, M. Westmoreland, and W. Wootters. Classical information capacity of a quantum channel. *Phys. Rev. A*, **54**, 1869, 1996.
- [98] Ma. Heid and N. Lütkenhaus. Security of coherent-state quantum cryptography in the presence of Gaussian noise. *Physical Review A* **76**, 022313, 2007.
- [99] W. Heisenberg. Über den anschaulichen inhalt der quantentheoretischen kinematik und mechanik. *Zeitschrift für Physik* **43**, 172, 1927.
- [100] C. W. Helstrom. *Quantum Detection and Estimation Theory*. Academic Press, New York, 1976.
- [101] M. Hillery and L. Mlodinow. Interferometers and minimum-uncertainty states. *Phys. Rev. A* **48**, 1548, 1993.
- [102] M. Hillery and L. Mlodinow. Correlated input-port, matter-wave interferometer: Quantum-noise limits to the atom-laser gyroscope. *Phys. Rev. A* **57**, 4736, 1998.
- [103] A. S. Holevo. *Probabilistic and Statistical Aspects of Quantum Theory*. North-Holland, Amsterdam, 1982.
- [104] A. S. Holevo. The capacity of the quantum channel with general signal states. *IEEE Trans. Inf. Theory* **44**, 269, 1998.
- [105] A. S. Holevo and N. R. F. Werner. Evaluationg capacities of bosonic Gaussian channels. *Phys. Rev. A* **63**, 032312, 2001.
- [106] C. K. Hong, Z. Y. Ou, and L. Mandel. Measurement of subpicosecond time intervals between two photons by interference. *Phys. Rev. Lett.* **59**, 2044, 1987.

- [107] K. Horodecki, M. Horodecki, P. Horodecki, and J. Oppenheim. Secure key from bound entanglement. *Phys. Rev. Lett.* **94**, 160502, 2005.
- [108] M. Horodecki, J. Oppenheim, and A. Winter. Partial quantum information. *Nature* **436**, 673, 2004.
- [109] M. Horodecki, J. Oppenheim, and A. Winter. Quantum state merging and negative information. *Comm. Math. Phys.* **269**, 107, 2006.
- [110] V. Jacques, E. Wu, F. Grosshans, F. Treussart, P. Grangier, A. Aspect, and J.-F. Roch. Experimental realization of Wheeler’s delayed-choice gedanken experiment. *Science* **315**, 966, 2007.
- [111] H. Jeong and M. S. Kim. Efficient quantum computation using coherent states. *Phys. Rev. A* **65**, 042305, 2002.
- [112] R. Jozsa and B. Schumacher. A new proof of the quantum noiseless coding theorem. *J Modern Optics* **41**, 2343, 1994.
- [113] L. Mišta Jr., Jr. R. Filip, and J. Fiurášek. Continuous-variable Werner state: Separability, nonlocality, squeezing, and teleportation. *Phys. Rev. A* **65**, 062315, 2002.
- [114] B. Julsgaard, J. Sherson, J. I. Cirac, J. Fiurášek, and E. S. Polzik. Experimental demonstration of quantum memory for light. *Nature* **432**, 482, 2004.
- [115] M.S. Kim, E. Park, P.L. Knight, and H. Jeong. Nonclassicality of a photon-subtracted gaussian field. *Phys. Rev. A* **71**, 043805, 2005.
- [116] Y. H Kim, R. Yu, S. P. Kulik, Yanhua Shih, and M. O. Scully. Delayed “choice” quantum eraser. *Phys. Rev. Lett.* **84**, 1, 2000.
- [117] C. King. Additivity for unital channels. *J. Math. Phys.* **43**, 4641, 2002.
- [118] A. Kitagawa, M. Takeoka, K. Wakui, and M. Sasaki. Effective squeezing enhancement via measurement-induced non-gaussian operation and its application to dense coding scheme. *Phys. Rev. A* **72**, 022334, 2005.
- [119] O. Klein. *Z. Phys.* **72**, 767, 1931.
- [120] M. Koniorczyk, Z. Kurucz, A. Gábris, and J. Janszky. General optical state truncation and its teleportation. *Phys. Rev. A* **62**, 013802, 2000.
- [121] P. G. Kwiat, P. H. Eberhard, A. M. Steinberg, and R. Y. Chiao. Proposal for a loophole-free Bell inequality experiment. *Phys. Rev. A* **49**, 3209, 1994.
- [122] P. G. Kwiat, K. Mattle, H. Weinfurter, A. Zeilinger, A. V. Sergienko, and Y. Shih. New high-intensity source of polarization-entangled photon pairs. *Phys. Rev. Lett.* **75**, 4337, 1995.
- [123] A. M. Lance, T. Symul, V. Sharma, Ch. Weedbrook, T. C. Ralph, and Ping Koy Lam. No-switching quantum key distribution using broadband modulated coherent light. *Phys. Rev. Lett.* **95**, 180503, 2005.
- [124] R. Landauer. Irreversibility and heat generation in the computing process. *IBM Journal of Research and Development*, vol. **5**, 183, 1961.
- [125] H. Lee, P. Kok, N.J. Cerf, and J.P. Dowling. Linear optics and projective measurements alone suffice to create large-photon-number path entanglement. *Phys. Rev. A* **65**, 030101(R), 2002.

- [126] J. A. Levenson, I. Abram, T. Rivera, P. Fayolle, J. C. Garreau, and P. Grangier. Quantum optical cloning amplifier. *Phys. Rev. Lett.* **70**, 267, 1993.
- [127] J. Lodewyck, T. Debuisschert, R. Tualle-Brouiri, and P. Grangier. Controlling excess noise in fiber-optics continuous-variable quantum key distribution. *Phys. Rev. A* **72**, 050303(R), 2005.
- [128] Jerome Lodewyck, Matthieu Bloch, Raul Garcia-Patron, Simon Fossier, Evgueni Karpov, Eleni Diamanti, Thierry Debuisschert, Nicolas J. Cerf, Rosa Tualle-Brouiri, Steven W. McLaughlin, and Philippe Grangier. Quantum key distribution over 25 km with an all-fiber continuous-variable system. *Phys. Rev. A (in press)*, 2007.
- [129] A.P. Lund, H. Jeong, T.C. Ralph, and M. S. Kim. Conditional production of superpositions of coherent states with inefficient photon detection. *Phys. Rev. A* **70**, 020101(R), 2004.
- [130] A. I. Lvovsky, H. Hansen, T. Aichele, O. Benson, J. Mlynek, and S. Schiller. Quantum state reconstruction of the single-photon fock state. *Phys. Rev. Lett.* **87**, 050402, 2001.
- [131] D.N. Matsukevich and A. Kuzmich. Quantum state transfer between matter and light. *Science* **306**, 663, 2004.
- [132] M. W. Mitchell, J. S. Lundeen, and A. M. Steinberg. Super-resolving phase measurements with a multiphoton entangled state. *Nature (London)* **429**, 161, 2004.
- [133] A. M. Mood, F. A. Graybill, and D. C. Boes. *Introduction to the Theory of Statistics*. Mc Graw-Hill, 1974.
- [134] W. J. Munro. Optimal states for Bell-inequality violations using quadrature-phase homodyne measurements. *Phys. Rev. A* **59**, 4197, 1999.
- [135] H. Nha and H. J. Carmichael. Proposed test of quantum nonlocality for continuous variables. *Phys. Rev. Lett.* **93**, 020401, 2004.
- [136] M.A. Nielsen and I.C. Chuang. *Quantum Computation and Quantum Information*. CUP, Cambridge, 2002.
- [137] C. Olalla. *La fuerza del deber, Planck*. Nivola, Madrid, 2006.
- [138] S. Olivares, M.G.A. Paris, and R. Bonifacio. Teleportation improvement by inconclusive photon subtraction. *Phys. Rev. A* **67**, 032314, 2002.
- [139] T. Opatrný, G. Kurizki, and D.-G. Welsch. Improvement on teleportation of continuous variables by photon subtraction via conditional measurement. *Phys. Rev. A* **61**, 032302, 2000.
- [140] Z. Y. Ou, S. F. Pereira, H. J. Kimble, and K. C. Peng. Realization of the Einstein-Podolsky-Rosen paradox for continuous variables. *Phys. Rev. Lett.* **68**, 3663, 1992.
- [141] A. Ourjoumtsev, H. Jeong, R. Tualle-Brouiri, and P. Grangier. Generation of optical 'Schrödinger cats' from photon number states. *Nature* **448**, 784, 2006.
- [142] A. Ourjoumtsev, R. Tualle-Brouiri, J. Laurat, and P. Grangier. Generating optical Schrödinger kittens for quantum information processing. *Science* **312**, 83, 2006.

- [143] J.-W. Pan, D. Bouwmeester, M. Daniell, H. Weinfurter, and A. Zeilinger. Experimental test of quantum nonlocality in three-photon greenberger–horne–zeilinger entanglement. *Nature* **403**, 515, 2000.
- [144] J.-W. Pan, M. Daniell, S. Gasparoni, G. Weihs, and A. Zeilinger. Experimental demonstration of four-photon entanglement and high-fidelity teleportation. *Phys. Rev. Lett.* **86**, 4435, 2001.
- [145] M.G.A. Paris. Displacement operator by beam splitter. *Phys. Lett. A* **217**, 78, 1996.
- [146] P. M. Pearle. Hidden-variable example based upon data rejection. *Phys. Rev. D* **2**, 1418, 1970.
- [147] D. T. Pegg, L. S. Phillips, and S. M. Barnett. Optical state truncation by projection synthesis. *Phys. Rev. Lett.* **81**, 1604, 1998.
- [148] A. Peres. *Quantum Theory: Concepts and Methods*. Kluwer Academic, Dordrecht, 1993.
- [149] M. J. Perrin. New experiments on the cathode rays. *Nature* **53**, 298, 1896.
- [150] M. Planck. On the law of distribution of energy in the normal spectrum. *Ann. D. Phys.* **4**, 553, 1901.
- [151] J. P. Poizat and P. Grangier. Experimental realization of a quantum optical tap. *Phys. Rev. Lett.* **70**, 271, 1993.
- [152] E. S. Polzik, J. Carri, and H. J. Kimble. Spectroscopy with squeezed light. *Phys. Rev. Lett.* **68**, 3020, 1992.
- [153] T.C. Ralph, A. Gilchrist, G.J. Milburn, W.J. Munro, and S. Glancy. Quantum computation with optical coherent states. *Phys. Rev. A* **68**, 042319, 2003.
- [154] J. M. Renes and G. Smith. Noisy processing nad distillation of private quantum states.
- [155] R. Renner. *Security of Quantum Key Distribution*. ETH Zürich, 2005.
- [156] R. Renner. Private communication. 2007.
- [157] R. Renner. Symmetry implies independence. *arXiv:quant-ph/0703069*, 2007.
- [158] R. Renner, N. Gisin, and B. Kraus. *Information-theoretic security proof for quantum-key-distribution protocols*. *Phys. Rev. A* **72**, 012332, 2005.
- [159] K. J. Resch, J. S. Lundeen, and A. M. Steinberg. Quantum state preparation and conditional coherence. *Phys. Rev. Lett.* **88**, 113601, 2002.
- [160] R. Rivest and L. Adleman A. Shamir. A method for obtaining digital signatures and public-key cryptosystems. *Communication of the ACM, Vol. 21 (2)*, pp. 120-126, 1978.
- [161] M.A. Rowe, D. Kielpinski, V. Meyer, C.A. Sackett, W. M. Itano, C. Monroe, and D.J. Wineland. Experimental violation of a Bell’s inequality with efficient detection. *Nature (London)* **409**, 791, 2001.
- [162] E. Rutherford. The scattering of α and β particles by matter and the structure of the atom. *Philosophical Magazine, Series 6 21*: 669–688, 1911.

- [163] E. Santos. Critical analysis of the empirical tests of local hidden-variable theories. *Phys. Rev. A* **46**, 3646, 1992.
- [164] E. Santos. Constraints for the violation of the Bell inequality in Einstein-Podolsky-Rosen-Bohm experiments. *Phys. Lett. A* **200**, 1, 1995.
- [165] V. Scarani, S. Iblisdir, N. Gisin, and A. Acin. Quantum cloning. *Rev. Mod. Phys.* **77**, 1225-1256, 2005.
- [166] E. Schmidt. Zur theorie der linearen und nichtlinearen integralgleichungen. *Math. Annalen.* **63**, 433, 1906.
- [167] C. Schori, J. L. Sørensen, and E. S. Polzik. Narrow-band frequency tunable light source of continuous quadrature entanglement. *Phys. Rev. A* **66**, 033802, 2002.
- [168] E. Schrödinger. An undulatory theory of the mechanics of atoms and molecules. *Phys. Rev.* **28**, 1049, 1926.
- [169] E. Schrödinger. Die gegenwärtige situation in der quantenmechanik. *Naturwissenschaften* **23**, 807, 1935.
- [170] B. Schumacher. Quantum coding. *Phys. Rev. A* **51**, 2738, 1995.
- [171] B. Schumacher and M. D. Westmoreland. Sending classical information via noisy quantum channels. *Phys. Rev. A* **56**, 131, 1997.
- [172] M. O. Scully and M. S. Zubairy. *Quantum Optics*. CUP, Cambridge, 1999.
- [173] A. Serafini. Detecting multimode entanglement by symplectic uncertainty relations. *quant-ph/0508231*, 2005.
- [174] A. Serafini. Multimode uncertainty relations and separability of continuous variable states. *Phys. Rev. Lett.* **96**, 110402, 2006.
- [175] C. E. Shannon. A mathematical theory of communication. *Bell System Technical Journal*, vol. **27**, pp. 379-423 and 623-656, 1948.
- [176] C. E. Shannon. Communication theory of secrecy systems. *Bell System Technical Journal*, Vol **28**, pp. 656-715, 1949.
- [177] P. Shor. Additivity of the classical capacity of entanglement breaking quantum channels. *J. Math. Phys.* **43**, 4334, 2002.
- [178] P. Shor. Equivalence of additivity questions in quantum information theory. *arXiv:quant-ph/0305035*, 2003.
- [179] P. Shor and J. Smolin. Quantum error-correcting codes need not completely reveal the error syndrome. *arXiv:quant-ph/9604006*, 1996.
- [180] P. W. Shor and J. Preskill. Simple proof of security of the BB84 quantum key distribution protocol. *Phys. Rev. Lett.* **85**, 441, 2000.
- [181] Ch. Silberhorn, T. C. Ralph, N. Lütkenhaus, and G. Leuchs. Continuous variable quantum cryptography: Beating the 3 db loss limit. *Phys. Rev. Lett.* **89**, 167901, 2002.
- [182] C. Simon and W.T.M. Irvine. Robust long-distance entanglement and a loophole-free Bell test with ions and photons. *Phys. Rev. Lett.* **91**, 110405, 2003.
- [183] R. Simon, S. Chaturvedi, and V. Srinivassan. *J. Math. Phys.* **40**, 3632, 1999.

- [184] S. Sing. *The Code Book: The Science of Secrecy from Ancient Egypt to Quantum Cryptography*. Anchor Books, 2000.
- [185] G. Smith and J. Smolin. Degenerate quantum codes for Pauli channels. *Phys. Rev. Lett.* **98**, 030501, 2007.
- [186] J. J. Thomson. Cathode rays. *Nature* **55**, 453, 1897.
- [187] J. J. Thomson. On computable numbers, with an application to the entscheidungsproblem. *Proc. Lon. Math. Soc., Series 2*, 42, 1936.
- [188] W. Tittel, J. Brendel, B. Gisin, T. Herzog, H. Zbinden, and N. Gisin. Experimental demonstration of quantum correlations over more than 10 km. *Phys. Rev. A* **57**, 3229, 1998.
- [189] A. Uhlmann. The 'transition probability' in the state space of a *-algebra. *Rep. Math. Phys.* **9**, 273, 1976.
- [190] G. van Assche. *Quantum Cryptography and Secret-Key Distillation*. CUP, Cambridge, 2006.
- [191] C. J. Villas-Boas, Y. Guimaraes, M.H.Y. Moussa, and B. Baseia. Recurrence formula for generalized optical state truncation by projection synthesis. *Phys. Rev. A* **63**, 055801, 2001.
- [192] J. von Neumann. *The Mathematical Foundations of Quantum Mechanics*. Princeton University Press, Princeton, 1996.
- [193] P. Walther, K.J. Resch, T. Rudolph, E. Schenck, H. Weinfurter, V. Vedral, M. Aspelmeyer, and A. Zeilinger. Experimental one-way quantum computing. *Nature (London)* **434**, 169, 2005.
- [194] C. Weedbrook, A. M. Lance, W. P. Bowen, T. Symul, T. C. Ralph, and P. K. Lam. Quantum cryptography without switching. *Phys. Rev. Lett.* **93**, 170504, 2004.
- [195] G. Weihs, T. Jennewein, C. Simon, H. Weinfurter, and A. Zeilinger. Violation of Bell's inequality under strict einstein locality conditions. *Phys. Rev. Lett.* **81**, 5039, 1998.
- [196] J. Wenger, M. Hafezi, F. Grosshans, R. Tualle-Brouri, and P. Grangier. Maximal violation of Bell inequalities using continuous-variable measurements. *Phys. Rev. A* **67**, 012105, 2003.
- [197] J. Wenger, R. Tualle-Brouri, and Ph. Grangier. Non-Gaussian statistics from individual pulses of squeezed light. *Phys. Rev. Lett.* **92**, 153601, 2004.
- [198] M. M. Wolf, G. Giedke, and J. I. Cirac. Extremality of Gaussian quantum states. *Phys. Rev. Lett.* **96**, 080502, 2006.
- [199] W. K. Wootters and W. H. Zurek. A single quantum cannot be cloned. *Nature* **299**, 802, 1982.
- [200] B. Yurke. Input states for enhancement of fermion interferometer sensitivity. *Phys. Rev. Lett.* **56**, 1515, 1986.
- [201] B. Yurke, S.L. McCall, and J.R. Klauder. SU(2) and SU(1,1) interferometers. *Phys. Rev. A* **33**, 4033, 1986.

- [202] A. Zavatta, S. Viciani, and M. Bellini. Tomographic reconstruction of the single-photon fock state by high-frequency homodyne detection. *Phys. Rev. A* **70**, 053821, 2004.
- [203] T. C. Zhang, K. W. Goh, C. W. Chou, P. Lodahl, and H. J. Kimble. Quantum teleportation of light beams. *Phys. Rev. A* **67**, 033802, 2003.
- [204] X.B. Zou, K. Pahlke, and W. Mathis. Phase measurement and generation of arbitrary superposition of fock states. *Phys. Lett. A* **323**, 329, 2004.